SECOND-ORDER REFINEMENTS FOR t-RATIOS WITH MANY INSTRUMENTS

YUKITOSHI MATSUSHITA AND TAISUKE OTSU

Abstract. This paper studies second-order properties of the many instruments robust t-ratios based on the limited information maximum likelihood and Fuller estimators for instrumental variable regression models with homoskedastic errors under the many instruments asymptotics, where the number of instruments may increase proportionally with the sample size n, and proposes second-order refinements to the t-ratios to improve the size and power properties. Based on asymptotic expansions of the null and non-null distributions of the t-ratios derived under the many instruments asymptotics, we show that the second-order terms of those expansions may have non-trivial impacts on the size as well as the power properties. Furthermore, we propose adjusted t-ratios whose approximation errors for the null rejection probabilities are of order $O(n^{-1})$ in contrast to the ones for the unadjusted t-ratios of order $O(n^{-1/2})$, and show that these adjustments induce some desirable power properties in terms of the local maximinity. Although these results are derived under homoskedastic errors, we also establish a stochastic expansion for a heteroskedasticity robust t-ratio, and propose an analogous adjustment under slight deviations from homoskedasticity.

1. Introduction

Instrumental variable regression is one of the most widely used methods in empirical economic analysis. Particularly in microeconometric applications, researchers often use many instrumental variables to improve efficiency of estimators and associated inference methods (e.g., Angrist and Krueger, 1991). However, in such cases, it has been found that approximate distributions of the estimators and statistics based on the conventional asymptotic theory can be inaccurate. For example, the two stage least squares (TSLS) estimator tends to have large bias. Although the limited information maximum likelihood (LIML) estimator is less biased, its distribution is often more dispersed than the limiting distribution based on the conventional asymptotics (see, e.g., Anderson, Kunitomo and Sawa, 1982, and Anderson, Kunitomo and Matsushita, 2010, 2011).

In order to give more accurate approximations under many instruments, Kunitomo (1980, 1982) and Morimune (1983) considered a limiting sequence where the number of instruments K is allowed to increase proportionally with the sample size n (called the large-K asymptotics), and derived the limiting distribution of the LIML estimator when the disturbances are normal and there is one endogenous regressor in the regression model. Bekker (1994) derived multivariate first-order approximations to the distributions of several estimators under the large-K asymptotics with the normal disturbances, while Hansen, Hausman and Newey (2008), van Hasselt (2010), and Anderson, Kunitomo and Matsushita (2010) extended those results to non-normal cases. Hausman *et al.* (2012) considered a more general model, where the reduced form may

We are grateful to Naoto Kunitomo, anonymous referees, and an associate editor for helpful comments. Matsushita acknowledges financial support from the JSPS KAKENHI (18K01541).

be nonlinear and the disturbances may be heteroskedastic, and suggested to use t-ratios based on heteroskedasticity and many instrument robust versions of the LIML and Fuller (1977) estimators and their asymptotic variance estimators. It should be noted that existing works on the large-K asymptotics mostly focus on the first-order asymptotic properties of the estimators and test statistics and their higher-order properties are largely unexplored.[1]

This paper studies second-order properties of the t-ratios based on the LIML and Fuller estimators for instrumental variable regression models with homoskedastic errors under the large-K asymptotics, and proposes higher-order refinements to the t-ratios to improve the size and power properties. To explore the finite sample properties of the t-ratios with many instruments, asymptotic expansions of the null and non-null distributions of the large-K robust t-ratios associated with the LIML and Fuller estimators are derived under the large-K asymptotics. Moreover, to assess the effects of variance estimation, we derive asymptotic expansions of the LIML and Fuller estimators under the large-K asymptotics. Based on these asymptotic expansions, it is shown that the finite sample distributions of the large-K t-ratios can be quite different from those of the corresponding standardized estimators although they have the same asymptotic normal distribution under the large-K asymptotics. In fact, the absolute values of the second-order terms of the asymptotic expansions of the standardized LIML and Fuller estimators and their large-K t-ratios are the same but have opposite signs. Also the null distributions of the large-K t-ratios can be skewed and largely deviated from the standard normal distribution. For two-sided testing, although the second-order terms cancel out under the null hypothesis, we find that these second-order terms may have non-trivial impacts on the power properties. Based on these expansions, we propose adjusted t-ratios whose approximation errors for the null rejection probabilities are of order $O(n^{-1})$ in contrast to the ones for the unadjusted t-ratios of order $O(n^{-1/2})$. Furthermore, we show that these adjustments induce some desirable power properties in terms of the local maximinity. Although these results are derived under homoskedastic errors, we also establish a stochastic expansion for the heteroskedasticity robust t-ratio based the heteroskedasticity and many instruments robust versions of the LIML estimator by Hausman *et al.* (2012), and propose an analogous adjustment under slight deviations from homoskedasticity. These findings are illustrated by some simulation studies.

This paper also contributes to the literature of the asymptotic higher-order expansion approach, which has been developed extensively to investigate the finite sample properties of econometric methods (see, e.g., Rothenberg, 1984, and Ullah, 2004, for an overview). For simultaneous equation models, it has been used to give more accurate approximations to distributions of estimators and test statistics, or to compare their higher-order properties, see, Anderson (1974), Sargan (1975), Phillips (1977), Rothenberg (1988), Fujikoshi *et al.* (1982), Morimune (1989), to name a few. Our main contribution in this context is that, to the best of our knowledge, this is the first paper which investigates higher-order properties of testing methods under the large-K asymptotics.

[1]Exceptions are Kunitomo (1980) and Morimune (1983) who established asymptotic expansions for the distributions of the LIML and k-class estimators, respectively, for the case of the normal disturbances and one endogenous regressor.

The paper is organized as follows. In Section 2, we introduce our setup and estimators (Section 2.1) and define the the large-K robust t-ratios (Section 2.2). Section 3 presents our main results: asymptotic expansions of the large-K t-ratios under the null hypothesis (Section 3.1), asymptotic expansions under the local alternatives and adjusted t-ratios (Section 3.2), and a discussion for the case of heteroskedastic disturbances (Section 3.3). Section 4 illustrates our findings by some simulation results.

2. Setup, estimators and test statistics

2.1. **Setup and estimators.** We first introduce our basic setup and some estimators. Consider a single structural equation

$$y_1 = Y_2\beta + Z_1\gamma + u, \tag{1}$$

where y_1 is an n-vector of dependent variables, Y_2 is an $n \times G_2$ matrix of endogenous regressors, Z_1 is an $n \times K_1$ matrix of exogenous regressors, u is an n-vector of error terms, and β and γ are G_2 and K_1 dimensional vectors of unknown parameters, respectively. We assume that (1) is the first equation in a simultaneous system of $G_2 + 1$ linear stochastic equations relating $G_2 + 1$ endogenous variables $Y = (y_1, Y_2)$, and $K_1 + K_2$ exogenous variables $Z = (Z_1, Z_2)$, where Z_2 is an $n \times K_2$ matrix of instrumental variables for (1). Let $K = K_1 + K_2$. The reduced form of Y is defined as

$$Y = Z\Pi + V = (Z_1, Z_2)\begin{pmatrix} \Pi_1 \\ \Pi_2 \end{pmatrix} + (v_1, V_2), \tag{2}$$

where $\Pi_1 = (\pi_{11}, \Pi_{12})$ and $\Pi_2 = (\pi_{21}, \Pi_{22})$ are $K_1 \times (1 + G_2)$ and $K_2 \times (1 + G_2)$ matrices, respectively, of the reduced form coefficients, and (v_1, V_2) is an $n\times(1+G_2)$ matrix of disturbances.

We impose the following assumptions on the reduced form.

Assumption 1. *Z is nonrandom. Π_{22} is of rank G_2. The rows of V are independently distributed, and each row of V has mean* 0 *and nonsingular covariance matrix* $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{pmatrix}$.

As in Hansen, Hausman and Newey (2008) and Hausman *et al.* (2012), we consider the case of nonrandom Z. We could allow Z to be random by modifying our assumptions, expectation, and probability to be conditional on Z as in Chao *et al.* (2012). The last assumption implies that the error term u is homoskedastic. We discuss the case of heteroskedastic errors in Section 3.3.

In order to relate (1) and (2), we postmultiply (2) by $(1, -\beta')'$. Then it can be written as in (1) with $u = v_1 - V_2\beta$, where the components of u are independently distributed with mean 0 and variance $\sigma^2 = \omega_{11} - 2\beta'\omega_{21} + \beta'\Omega_{22}\beta$. Also, it holds $\gamma = \pi_{11} - \Pi_{12}\beta$ and $\pi_{21} = \Pi_{22}\beta$.

Let $P = Z(Z'Z)^{-1}Z'$ and $M = I - P$. The k-class estimator is defined as

$$\begin{pmatrix} \hat\beta_k \\ \hat\gamma_k \end{pmatrix} = \begin{bmatrix} Y_2'Y_2 - kY_2'MY_2 & Y_2'Z_1 \\ Z_1'Y_2 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{pmatrix} Y_2'(I - kM)y_1 \\ Z_1'y_1 \end{pmatrix}. \tag{3}$$

This estimator covers (i) OLS ($k = 0$), (ii) TSLS ($k = 1$), (iii) LIML ($k = \hat{\lambda}$), and (iv) Fuller (1977) ($k = \hat{\lambda} - a/(n - K)$ for some $a > 0$) as special cases, where $\hat{\lambda}$ is the smallest root of

$$\left| \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} P(Y, Z_1) - \hat{\lambda} \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} M(Y, Z_1) \right| = 0.$$

Under the conventional asymptotics, where the number of instruments K is fixed, both the LIML and TSLS estimators are consistent and follow the same limiting normal distribution. However, it has been known that the exact distributions of these estimators can be quite different from the normal distribution. When K is large, the TSLS estimator can be severely biased (see, e.g., Anderson, Kunitomo and Matsushita, 2010). On the other hand, the distribution of the LIML estimator is more dispersed than the limiting distribution under the conventional asymptotics. The Fuller estimator has moments of all orders, and is known to have good finite sample properties in some situations (see, e.g., Hahn, Hausman and Kuersteiner, 2004, and Hansen, Hausman and Newey, 2008).

Bekker (1994) pointed out that the large-K asymptotic theory, where K may grow proportionally to n, may be suited better to applications, even when the number of instruments is not large. Under the large-K asymptotics, the (first-order) asymptotic distributions of the LIML and TSLS estimators are rather different. The LIML estimator is consistent and asymptotically normal while the TSLS estimator loses consistency. Also the LIML estimator attains the asymptotic efficiency bound when the number of instruments is large (see, Kunitomo, 1982, Chioda and Jansson, 2009, and Anderson, Kunitomo and Matsushita, 2010).

In contrast, this paper is concerned with higher-order properties of the t-ratios for testing parameter hypotheses under the large-K asymptotics, which will be introduced in the next subsection.

2.2. **Many instruments robust t-statistics.** Let ι be a $(G_2 + K_1)$-vector of zeros, apart from its j-th element which is unity. We are interested in testing the null hypothesis

$$H_0 : \iota' \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = 0,$$

i.e., the j-th coefficient in (1) is zero, against the one-sided or two-sided alternative hypothesis.

The focus of this paper is to investigate higher-order properties of the t-tests for H_0 under the large-K asymptotics. In particular, we impose the following assumptions.

Assumption 2.

(i): *$K = K_n$ satisfies*

$$\frac{K}{n} = c + O(n^{-1}) \quad \textit{for some } c \in [0, 1). \tag{4}$$

(ii): *The distribution of the rows of V have finite 8th moments and belong to the class of elliptically contoured distributions, i.e., the characteristic function of the rows of V (say, v_i) has the form of $E[e^{\mathrm{i}t'v_i}] = \varphi(t'\Upsilon t)$ for some $\varphi(\cdot)$ and positive definite Υ, where* $\mathrm{i} = \sqrt{-1}$. *Furthermore,* $\max_{1 \le i \le n} ||D_2' z_i||^2 = o(n)$.

(iii): *For some positive definite matrix Q, it holds*

$$\frac{1}{n}D_2'Z'ZD_2 = Q + O(n^{-1}), \tag{5}$$

where $D_2 = \begin{pmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{pmatrix}$.

Assumption 2 (i) says that the number of instruments K can grow at either the same rate ($c > 0$) as the sample size or at slower rate ($c = 0$), where the latter includes the conventional fixed-K asymptotics as a special case. Assumption 2 (ii) greatly simplifies our higher-order analysis in the following section. See Anderson (2003, Section 2.7) for a discussion on elliptically contoured distributions. Assumption 2 (iii), also imposed in e.g., Kunitomo (1980) and Morimune (1989), is necessary for our asymptotic expansion.

Under Assumptions 1 and 2, the results in Anderson, Kunitomo and Matsushita (2010) can be adapted to derive the limiting distributions of the LIML estimator $(\hat{\beta}_{LI}', \hat{\gamma}_{LI}')'$ and Fuller estimator $(\hat{\beta}_F', \hat{\gamma}_F')'$ as

$$\sqrt{n}\begin{pmatrix} \hat{\beta}_{LI} - \beta \\ \hat{\gamma}_{LI} - \gamma \end{pmatrix} \xrightarrow{d} N(0, \Psi^*), \qquad \sqrt{n}\begin{pmatrix} \hat{\beta}_{F} - \beta \\ \hat{\gamma}_{F} - \gamma \end{pmatrix} \xrightarrow{d} N(0, \Psi^*),$$

where

$$\Psi^* = \sigma^2 Q^{-1} + \frac{c}{1-c}Q^{-1}\left[\begin{pmatrix} \Omega_{22}\sigma^2 & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4\right]Q^{-1} + Q^{-1}[(\Xi_3 + \Xi_3') + \eta \Xi_4]Q^{-1}. \tag{6}$$

For Ψ^*, we use the following notation

$$\begin{aligned} q_2 &= \frac{1}{\sigma^2}(\omega_{21}' - \beta'\Omega_{22}, 0')', \qquad \Xi_3 = \frac{1}{1-c}\lim_{n\to\infty} D_2'\frac{1}{n}\sum_{i=1}^{n} z_i (P_{ii} - c)E[u_i^2 w_{2i}'], \\ \eta &= \frac{1}{(1-c)^2}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(P_{ii} - c)^2, \qquad \Xi_4 = E[u_i^2 w_{2i} w_{2i}'] - \sigma^2 E[w_{2i} w_{2i}'], \end{aligned}$$

where $w_{2i} = (v_{2i}', 0')' - u_i q_2$ and $P_{ii} = z_i'(Z'Z)^{-1}z_i$. Compared to the conventional variance formula $\sigma^2 Q^{-1}$ under the fixed-K asymptotics, there are two additional terms in Ψ^* due to the large-K asymptotics, which vanish when the number of instruments grows at a slower rate (i.e., $c = 0$).

Under Assumption 2 (iii), the asymptotic variance Ψ^* simplifies to (see, Anderson, Kunitomo and Matsushita, 2010)

$$\Psi = \sigma^2 Q^{-1} + \left(\frac{c}{1-c} + \eta\kappa\right)Q^{-1}\left[\begin{pmatrix} \Omega_{22}\sigma^2 & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2' \sigma^4\right]Q^{-1}, \tag{7}$$

where $\kappa = (E[u_i^4]/\sigma^4 - 3)/3$. Note that Ψ is identical to Bekker's (1994) variance when the error terms are normally distributed. By taking the sample counterparts, a consistent estimator of Ψ

is given by

$$
\begin{aligned}
\hat{\Psi} &= \hat{\sigma}^2 \hat{Q}^{-1} \\
&+ \left(\frac{K}{n-K} + \hat{\eta}\hat{\kappa} \right) \hat{Q}^{-1} \begin{pmatrix} \frac{1}{n-K} Y_2' M Y_2 \hat{\sigma}^2 - \frac{1}{(n-K)^2} Y_2' M Y \hat{b}\hat{b}' Y' M Y_2 & 0 \\ 0 & 0 \end{pmatrix} \hat{Q}^{-1}, \quad (8)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-K} \hat{b}' Y' M Y \hat{b}, \qquad \hat{b} = (1, -\hat{\beta}')', \qquad \hat{Q} = \frac{1}{n} \begin{pmatrix} Y_2' P Y_2 - \lambda Y_2' M Y_2 & Y_2' Z_1 \\ Z_1' Y_2 & Z_1' Z_1 \end{pmatrix}, \\
\hat{\eta} &= \left(\frac{n}{n-K} \right)^2 \frac{1}{n} \sum_{i=1}^{n} \left(P_{ii} - \frac{K}{n} \right)^2, \qquad \hat{\kappa} = \frac{1}{3} \left\{ \frac{1}{\hat{\sigma}^4} \frac{1}{n} \sum_{i=1}^{n} (y_{1i} - y_{2i}' \hat{\beta} - z_{1i}' \hat{\gamma})^4 - 3 \right\},
\end{aligned}
$$

and $(\hat{\beta}, \hat{\gamma}, \lambda) = (\hat{\beta}_{LI}, \hat{\gamma}_{LI}, \hat{\lambda})$ for the LIML estimator, or $(\hat{\beta}_F, \hat{\gamma}_F, \hat{\lambda} - a/(n-K))$ for the Fuller estimator.

Based on this variance estimator, the large-K t-ratio for testing H_0 is given by

$$
t_K = \frac{1}{\sqrt{\hat{\Psi}_j}} \iota' \sqrt{n} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}, \tag{9}
$$

where $\hat{\Psi}_j$ is the j-th diagonal element of $\hat{\Psi}$. In the next section, we study higher-order properties of the t-test based on this statistic under the large-K asymptotics.

3. Main results

3.1. **Asymptotic expansions under null hypothesis.** We first present an asymptotic expansion of the null distribution of the t-ratio in (9) under the large-K asymptotics in (4). Let $\Phi(\cdot)$ and $\phi(\cdot)$ be the standard normal distribution and density functions, respectively, and Ψ_j be the j-th diagonal element of Ψ in (7).

Theorem 1. *Suppose Assumptions 1 and 2 hold true. Then the asymptotic expansion of the null distribution of the large-K t-ratio t_K in (9) using the Fuller estimator $(\hat{\beta}, \hat{\gamma}, \lambda) = (\hat{\beta}_F, \hat{\gamma}_F, \hat{\lambda} - a/(n-K))$ is given by*

$$
P\{t_K \le \tau\} = \Phi(\tau) - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\Psi_j}} \left[\tau^2 \iota' \Psi + \frac{a}{1-c} \sigma^2 \iota' Q^{-1} \right] q_2 \phi(\tau) + O(n^{-1}), \tag{10}
$$

for each $\tau \in \mathbb{R}$. Furthermore, the expansion for t_K using the LIML estimator $(\hat{\beta}, \hat{\gamma}, \lambda) = (\hat{\beta}_{LI}, \hat{\gamma}_{LI}, \hat{\lambda})$ is obtained by setting $a = 0$ in (10).

This theorem says that (i) the approximation errors in rejection probability of both tests are of order $O(n^{-1/2})$, (ii) the approximation errors become larger as the degree of endogeneity $q_2 = E[v_{2i} u_i]/\sigma^2$ increases, and (iii) the approximation error of t_K by the Fuller estimator is not always smaller than that of t_K by the LIML estimator. In particular, if the number of endogenous regressors is one (i.e., $G_2 = 1$), the absolute value of the second-order term in (10) by the Fuller estimator is always larger than the one by the LIML estimator unless $q_2 = 0$. This point implies that improvement of the Fuller estimator over the LIML estimator does not necessarily imply improvement of the size property of the t-test under the large-K asymptotics.

It should be noted that the finite sample distributions of the t-ratios and corresponding estimators may be different due to estimation of the asymptotic variances in the denominators of t-ratios. In fact, we can derive the asymptotic expansions of the distributions of the LIML and Fuller estimators as follows. Let ϕ_Ψ be the density function of $N(0,\Psi)$.

Theorem 2. *Suppose Assumptions 1 and 2 hold true. Then the asymptotic expansion of the density f of $\sqrt{n}((\hat{\beta}-\beta)',(\hat{\gamma}-\gamma)')'$ using the Fuller estimator $(\hat{\beta},\hat{\gamma},\lambda)=(\hat{\beta}_F,\hat{\gamma}_F,\hat{\lambda}-a/(n-K))$ is given by*

$$f(\xi)=\phi_\Psi(\xi)\left\{1+\frac{1}{\sqrt{n}}\left[(q_2'\xi)(G_1+K_1+1-\xi'\Psi\xi)+\frac{a}{1-c}q_2'\sigma^2Q^{-1}\Psi\xi\right]\right\}+O(n^{-1}), \qquad (11)$$

for each $\xi\in\mathbb{R}^{G_2+K_1}$. Furthermore, the expansion of the density f of $\sqrt{n}((\hat{\beta}-\beta)',(\hat{\gamma}-\gamma)')'$ using the LIML estimator $(\hat{\beta},\hat{\gamma},\lambda)=(\hat{\beta}_{LI},\hat{\gamma}_{LI},\hat{\lambda})$ is obtained by setting $a=0$ in (11).

We note that the expansion for the LIML estimator is derived based on Fujikoshi *et al.* (1982), but the expansion for the Fuller estimator is new in the literature. Based on Theorem 2, we can see that under H_0,

$$P\left\{\frac{1}{\sqrt{\Psi_j}}\iota'\sqrt{n}\begin{pmatrix}\hat{\beta}\\ \hat{\gamma}\end{pmatrix}\le\tau\right\}=\Phi(\tau)+\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_j}}\left[\tau^2\iota'\Psi+\frac{a}{1-c}\sigma^2\iota'Q^{-1}\right]q_2\phi(\tau)+O(n^{-1}),$$

for each $\tau\in\mathbb{R}$. Comparing this expansion with (10), we can see that the distributions of the LIML and Fuller estimators on the one hand and large-K t-ratio on the other are distorted in opposite directions.[2]

3.2. **Asymptotic expansions under local alternative hypothesis and adjusted t-ratios.** In order to investigate power properties of the large-K t-ratios, we now derive asymptotic expansions of their distributions under the local alternative hypothesis:

$$H_{1n}:\iota'\begin{pmatrix}\beta\\ \gamma\end{pmatrix}=\frac{1}{\sqrt{n}}\iota'\zeta,$$

for some $\zeta\in\mathbb{R}^{G_2+K_1}$. Let Φ_ζ and ϕ_ζ be the cumulative distribution and density functions of $N(\Psi_j^{-1/2}\iota'\zeta,1)$, respectively.

Theorem 3. *Suppose Assumptions 1 and 2 hold true. Then asymptotic expansions of the power functions of the one-sided and two-sided large-K t-tests under H_{1n} are given by*

$$\begin{aligned}P\{t_K\ge\tau\} &= 1-\Phi_\zeta(\tau)\\ &\quad+\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_j}}\left[(\iota'\Psi q_2)\left\{\tau^2-\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right\}+\frac{a\sigma^2}{1-c}\iota'Q^{-1}q_2\right]\phi_\zeta(\tau)+O(n^{-1}),\\ P\{|t_K|\ge\tau\} &= 1-\{\Phi_\zeta(\tau)-\Phi_\zeta(-\tau)\}\\ &\quad+\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_j}}\left[(\iota'\Psi q_2)\left\{\tau^2-\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right\}+\frac{a\sigma^2}{1-c}\iota'Q^{-1}q_2\right]\{\phi_\zeta(\tau)-\phi_\zeta(-\tau)\}+O(n^{-1}),\end{aligned}$$

[2]Bekker (1994) provided a skewed approximation for the distribution of the LIML estimator, not its t-ratio. Because the directions of the skewness of the distributions of the LIML estimator and the t-ratio under the null are opposite, Bekker's skewed approximation might make the size property of the test even worse.

for each $\tau \in \mathbb{R}$, respectively. The expansions for the t-ratios using the LIML estimator is obtained by setting $a = 0$.

This theorem says that the large-K t-tests are locally biased up to the order $O(n^{-1/2})$ unless the degree of endogeneity q_2 is zero because $\frac{\partial}{\partial \zeta} P\{t_K \geq \tau\} = O(n^{-1/2})$ and $\frac{\partial}{\partial \zeta} P\{|t_K| \geq \tau\} = O(n^{-1/2})$ unless $q_2 = 0$ (Rao, 1973, pp.454).

Based on the asymptotic expansions in Theorems 1 and 3, we propose a simple adjustment to the t-ratio, which does not include the term of order $O(n^{-1/2})$ in the expansion of the null distribution:

$$t_K^{adj} = t_K - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{\hat{\Psi}_j}} \left\{ \iota' \hat{\Psi} \hat{q}_2 t_{K0}^2 + \frac{an}{n-K} \hat{\sigma}^2 \iota' \hat{Q}^{-1} \hat{q}_2 \right\}, \tag{12}$$

where $\hat{q}_2 = \frac{1}{\hat{\sigma}^2} \frac{1}{n-K} (Y_2, 0)' M Y (1, -\hat{\beta}')'$ and

$$t_{K0} = \frac{1}{\sqrt{\hat{\Psi}_{0j}}} \iota' \sqrt{n} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}, \tag{13}$$

with $\hat{\Psi}_{0j}$, the j-th diagonal element of $\hat{\Psi}$ with $(\hat{\beta}, \hat{\gamma})$ replaced by the constrained LIML or Fuller estimator under H_0. We note that computation of the adjusted statistic t_K^{adj} does not involve the constant c in Assumption 2 (i), which is replaced with K/n in finite samples.

The asymptotic expansion of the null distribution of the adjusted statistic is obtained as follows.

Theorem 4. *Suppose Assumptions 1 and 2 hold true. Then the asymptotic expansion of the null distribution of the adjusted t-ratio in (12) is given by*

$$P\{t_K^{adj} \leq \tau\} = \Phi(\tau) + O(n^{-1}), \tag{14}$$

for each $\tau \in \mathbb{R}$.

This theorem says that the approximation error in rejection probability of the one-sided test using the adjusted t-ratio t_K^{adj} is of order $O(n^{-1})$, which improves the (unadjusted) large-K t-test whose error is $O(n^{-1/2})$. We note that this result holds true for any $a > 0$. Popular choices are $a = 0$ (LIML), 1 (Fuller, 1977), or $K_2 - G_2$ (Anderson, Kunitomo and Morimune, 1986). In order to study the effect of a on the null distribution of t_K^{adj}, we need a detailed analysis on the higher-order terms of order $O(n^{-1})$ in (14), which is beyond the scope of this paper.[3]

Next we present the local power properties of the adjusted t-tests.

[3]Morimune (1983) derived the large-K asymptotic expansions for the distributions of the k-class estimators up to the order $O(n^{-1})$ for the case of the normal disturbances and one endogenous regressor, and suggested a choice of a based on the expansion.

Theorem 5. *Suppose Assumptions 1 and 2 hold true. Then the asymptotic expansions of the power functions of the one-sided and two-sided adjusted t-tests under H_{1n} are given by*

$$\begin{aligned} P\{t_K^{adj} \geq \tau\} &= 1 - \Phi_\zeta(\tau) + \frac{1}{\sqrt{n}}\left[-\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right]\phi_\zeta(\tau) + O(n^{-1}), \\ P\{|t_K^{adj}| \geq \tau\} &= 1 - \{\Phi_\zeta(\tau) - \Phi_\zeta(-\tau)\} + \frac{1}{\sqrt{n}}\left[-\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right]\{\phi_\zeta(\tau) - \phi_\zeta(-\tau)\} + O(n^{-1}), \end{aligned}$$

for each $\tau \in \mathbb{R}$, respectively.

Similar to Theorem 4, we note that the terms of order $O(n^{-1/2})$ do not depend on a. This is due to the fact that the adjustment term in (12) eliminates dependence on a from the conditional mean of the second-order term of t_K^{adj} given the first-order component (see eq. (36) in Appendix). Indeed this elimination occurs both under H_0 and H_{1n}.

Note that the adjusted t-tests do not dominate the unadjusted t-tests in terms of the second-order power uniformly in ζ. However, compared to the two-sided t-test using the LIML estimator with $a = 0$ (denoted by t_K^{LI}), the two-sided adjusted t-test t_K^{adj} has some desirable power property with regard to the second-order local maximinity. More precisely, we obtain the following result.

Proposition 1. *Suppose Assumptions 1 and 2 hold true. Under H_{1n}, there exists $\Delta > 0$ such that*

$$\lim_{n\to\infty} \sqrt{n}\left[\min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} P\{|t_K^{adj}| \geq \tau\} - \min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} P\{|t_K| \geq \tau\}\right] > 0,$$

for all $\delta \in (0, \Delta)$.

This proposition says that the adjusted t-ratio t_K^{adj} for two-sided hypothesis testing is more powerful than the unadjusted one t_K on the basis of the minimum power attainable for alternatives within a given distance from the null hypothesis. This result is similar to Mukerjee (1992) for the comparison between the conditional and usual likelihood ratio statistics. Note that since $\sqrt{n}\left[P\{|t_K^{adj}| \geq \tau\} - P\{|t_K| \geq \tau\}\right] \to 0$ from Theorems 3 and 5 by setting $\zeta = 0$, the above power superiority of t_K^{adj} is not a consequence of size distortion.

3.3. **Discussion: Heteroskedastic errors.** This paper should be considered as a starting point toward more general higher-order theory for inference on instrumental variable regression models under the large-K asymptotics. In particular, it is interesting to extend our analysis for the case of heteroskedastic error terms. It should be noted that the LIML and Fuller estimators can be inconsistent with many instruments and heteroskedasticity of unknown form (see, e.g., Hausman *et al.,* 2012). Hausman *et al.* (2012) proposed heteroskedasticity robust versions of the LIML and Fuller estimators and associated t-tests based on their asymptotic variance estimators. In this subsection, we study higher-order properties of the heteroskedasticity and many instruments robust t-tests by Hausman *et al.* (2012) and propose an adjusted t-statistic, which exhibits desirable size properties under slight deviations from homoskedasticity.

Throughout this subsection, let $X = (Y_2, Z_1)$ and P^* be an $n \times n$ matrix such that $P^*_{ij} = P_{ij}$ for $i \neq j$ and $P_{ii} = 0$. The heteroskedasticity robust version of the LIML (called HLIM) estimator for $\theta = (\beta', \gamma')'$ by Hausman *et al.* (2012) is defined as

$$\hat{\theta}_{HLIM} = (X'P^*X - \hat{\alpha}X'X)^{-1}(X'P^*y_1 - \hat{\alpha}X'y_1), \tag{15}$$

where $\hat{\alpha}$ is the smallest root of

$$\left| \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} P^*(Y, Z_1) - \hat{\alpha} \begin{pmatrix} Y' \\ Z_1' \end{pmatrix} (Y, Z_1) \right| = 0.$$

Hausman *et al.* (2012) established consistency and asymptotic normality of this estimator under the large-K asymptotics allowing heteroskedastic errors. Hausman *et al.* (2012) also proposed a consistent estimator for the asymptotic variance of $\hat{\theta}_{HLIM}$ and the t-ratio for testing $H_0 : \iota'\theta = 0$, that is

$$t_K^H = \frac{\iota'\sqrt{n}\hat{\theta}_{HLIM}}{\sqrt{\hat{\Psi}_{H,j}}}, \tag{16}$$

where $\hat{\Psi}_{H,j}$ is the j-th diagonal component of $\hat{\Psi}_H = \hat{Q}_H^{-1}\hat{\Sigma}_H\hat{Q}_H^{-1}$ with

$$\begin{aligned} \hat{Q}_H &= \frac{1}{n}(X'P^*X - \hat{\alpha}X'X), \\ \hat{\Sigma}_H &= \frac{1}{n}\sum_{k=1}^n \sum_{i \neq k}^n \sum_{j \neq k}^n \hat{X}_i P_{ik} \hat{u}_k^2 P_{kj} \hat{X}_j' + \frac{1}{n}\sum_{i=1}^n \sum_{j \neq i}^n \hat{X}_i \hat{X}_j' \hat{u}_i \hat{u}_j P_{ij}^2, \end{aligned}$$

$\hat{u}_i = y_i - X_i'\hat{\theta}_{HLIM}$, $\hat{X}_i = X_i - \hat{q}_2\hat{u}_i$, and $\hat{q}_2 = X'\hat{u}/(\hat{u}'\hat{u})$.

By an analogous but more lengthy argument as the proof of Theorem 1, we can derive the following asymptotic expansion for the heteroskedasticity robust t-ratio t_K^H. To this end, we modify our assumptions as follows.

Assumption 1H. *Z is nonrandom. Π_{22} is of rank G_2. The rows of V are independently distributed, and the i-th row of V has mean 0 and nonsingular covariance matrix Ω_i.*

Assumption 2H. *Suppose Assumption 2 (i) and (ii) hold true. Also, for some positive definite matrix Q_H, it holds*

$$\frac{1}{n}\sum_{i=1}^n D_2' z_i z_i'(1 - P_{ii})D_2 = Q_H + O(n^{-1}).$$

These assumptions are adapted from Hausman *et al.* (2012) to our context. Assumptions 1H and 2H are analogous to Assumptions 1 and 2, respectively, except for heteroskedasticity in V. Nonrandomness of Z is also imposed in Hausman *et al.* (2012).

Under these assumptions, a stochastic expansion of the the heteroskedasticity robust t-ratio t_K^H in (16) is obtained as follows.

Proposition 2. *Suppose Assumptions 1H and 2H hold true. Then under H_0, it holds*

$$t_K^H = \mathcal{T}_H + \frac{1}{\sqrt{n}} t_H^{(1)} + O_p(n^{-1}), \tag{17}$$

where

$$\mathcal{T}_H = \frac{\iota' e_H^{(0)}}{\sqrt{\Psi_{H,j}}}, \qquad t_H^{(1)} = \frac{\iota' e_H^{(1)}}{\sqrt{\Psi_{H,j}}} - \frac{1}{2}\frac{\Psi_{H,j}^{(1)}}{\Psi_{H,j}}\mathcal{T}_H,$$

$\Psi_{H,j}$ is the j-th diagonal element of $\Psi_H = Q_H^{-1}\Sigma_H Q_H^{-1}$, $\Psi_{H,j}^{(1)}$ is the j-th diagonal element of $\Psi_H^{(1)} = -Q_H^{-1}Q_H^{(1)}\Psi_H - \Psi_H Q_H^{(1)} Q_H^{-1} + Q_H^{-1}\Sigma_H^{(1)} Q_H^{-1}$, $\Sigma_H^{(1)} = \sqrt{n}(\hat{\Sigma}_H - \Sigma_H)$, $\sigma_i^2 = E[u_i^2]$,

$$\begin{aligned}
\Sigma_H &= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{i,j,k,i\neq j,j\neq k}^{n}\sigma_k^2 D_2' z_i P_{ik} P_{kj} z_j' D_2 + \frac{1}{n}\sum_{i\neq j}^{n} P_{ij}^2\{E[w_{2i}w_{2i}']\sigma_j^2 + E[w_{2i}u_i]E[w_{2j}u_j]\}\right], \\
Q_H^{(1)} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} D_2' z_i (1-P_{ii})(v_{2i}', 0') + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(v_{2i}', 0')'(1-P_{ii}) z_i D_2 + \frac{1}{\sqrt{n}}\sum_{i\neq j}^{n} P_{ij}(v_{2i}', 0')'(v_{2j}', 0') \\
&\quad -\alpha^{(1)}\left\{\bar{Q} + \begin{pmatrix} \bar{\Omega}_{22} & 0 \\ 0 & 0 \end{pmatrix}\right\},
\end{aligned}$$

and $e_H^{(0)}$, $e_H^{(1)}$, and $\alpha^{(1)}$ are defined in Appendix A.6.

Note that the asymptotic expansion of t_K^H takes an analogous form as the one of t_K in eq. (28) in Appendix for homoskedastic errors. Motivated by this proposition, we propose the following adjusted statistic for heteroskedastic errors:

$$t_K^{H,adj} = t_K^H - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\hat{\Psi}_{H,j}}}\iota'\hat{\Psi}_H \hat{q}_2 (t_{K0}^H)^2, \tag{18}$$

where

$$t_{K0}^H = \frac{\iota'\sqrt{n}\hat{\theta}_{HLIM}}{\sqrt{\hat{\Psi}_{H0,j}}}, \tag{19}$$

and $\hat{\Psi}_{H0,j}$ is the j-th diagonal element of $\hat{\Psi}_{H0} = \hat{Q}_H^{-1}\hat{\Sigma}_{H0}\hat{Q}_H^{-1}$ with

$$\hat{\Sigma}_{H0} = \frac{1}{n}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}\hat{X}_i P_{ik}\hat{u}_{0k}^2 P_{kj}\hat{X}_j' + \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\hat{X}_i\hat{X}_j'\hat{u}_{0i}\hat{u}_{0j}P_{ij}^2,$$

$\hat{u}_{0i} = y_i - X_i'\tilde{\theta}_{HLIM}$ with the constrained HLIM estimator $\tilde{\theta}_{HLIM}$ under H_0, $\hat{X}_i = X_i - \hat{q}_{20}\hat{u}_{0i}$, and $\hat{q}_{20} = X'\hat{u}_0/(\hat{u}'\hat{u})$.

By restricting the effect of heteroskedasticity, we can obtain analogous results to Theorems 1 and 4 as follows. Let $\bar{q}_2 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E[(v_{2i}', 0)'u_i]/\bar{\sigma}^2$ with $\bar{\sigma}^2 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E[u_i^2]$.

Proposition 3. *Suppose Assumptions 1H and 2H hold true. Additionally assume $\frac{1}{n}\sum_{i=1}^{n}(\Omega_i - \bar{\Omega})\otimes(\Omega_i - \bar{\Omega}) \to 0$. Then under H_0,*

$$\begin{aligned}
P\{t_K^H \leq \tau\} &= \Phi(\tau) - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_{H,j}}}(\tau^2\iota'\Psi_H)\bar{q}_2\phi(\tau) + O(n^{-1}), \\
P\{t_K^{H,adj} \leq \tau\} &= \Phi(\tau) + O(n^{-1}).
\end{aligned}$$

This proposition says that under the slight deviations from homoskedasticity (in the sense of $\frac{1}{n}\sum_{i=1}^{n}(\Omega_i - \bar{\Omega})\otimes(\Omega_i - \bar{\Omega}) \to 0$), the adjusted statistic $t_K^{H,adj}$ shows better null rejection properties. Without this condition, many additional higher-order terms emerge and adjustment on the t-ratio will be substantially more complicated. Full investigation without this condition is left for future research.

4. Simulation

In this section, we conduct a simulation study to examine quality of the preceding asymptotic approximations to the finite sample distributions of the t-ratios. We consider the data generating process (DGP):

$$\begin{aligned} y_{1i} &= y_{2i}\beta_0 + z_{1i}\gamma_0 + u_i, \\ y_{2i} &= z_i'\pi_2 + v_{2i}, \end{aligned} \tag{20}$$

for $i = 1, \ldots, n$, where $\pi_2 = (d, \ldots, d)'$, $z_i = (z_{1i}, z_{2i}')'$, $z_{1i} = 1$, and $z_{2i} \sim N(0, I_{K-1})$. The error terms are generated as $(u_i, v_{2i}) = (\epsilon_{1i}, \rho\epsilon_{1i} + \sqrt{1-\rho^2}\epsilon_{2i})$, where ϵ_{1i} and ϵ_{2i} are independent and drawn from $N(0,1)$.[4] We set $\beta_0 = \gamma_0 = 0$ and $\rho = 0.4, 0.8$, and the sample size is set as $n = 200$ for all cases. For each Monte Carlo replication, we set the value of d to fix the value of the concentration parameter as $\mu^2 = 30, 60$ (given the realized values of $\{z_i\}$), where

$$\mu^2 = \frac{\pi_2' \left[\sum_{i=1}^n z_{2i}z_{2i}' - \sum_{i=1}^n z_{2i}z_{1i}' \left(\sum_{i=1}^n z_{1i}z_{1i}'\right)^{-1} \sum_{i=1}^n z_{1i}z_{2i}'\right] \pi_2}{Var(v_{2i})}. \tag{21}$$

Null distributions of t-ratios for β_0. First, we investigate the null distributions of the five types of t-ratios – the standard t-ratio with the LIML estimator ($t^{LI} = \sqrt{n}\iota'\hat{\beta}_{LI}/\sqrt{\hat{\sigma}^2(\hat{Q}^{-1})_j}$), the large-$K$ t-ratio with the LIML estimator (t_K^{LI}), the large-K t-ratio with the Fuller estimator (t_K^F), the adjusted large-K t-ratio with the LIML estimator ($t_K^{adj,LI}$), and the adjusted large-K t-ratio with Fuller estimator ($t_K^{adj,F}$). The number of Monte Carlo repetitions in each experiment is 50,000.

Tables 1 and 2 report the null rejection frequencies of the one-sided and two-sided tests at the nominal 5% significance level, respectively. Our findings are summarized as follows.

- **i):** The size distortions of the standard t-ratio (t^{LI}) tend to be large when μ^2 is small and K is large. As K increases, the tails of t^{LI} become thicker and its rejection frequencies tend to be larger than the nominal level.
- **ii):** Compared to t^{LI}, the rejection frequencies of the large-K t-ratios (t_K^{LI} and t_K^F) are smaller and avoid over-rejection. Although t_K^{LI} and t_K^F work well for two-sided testing, they show some asymmetric behaviors for one-sided testing. More precisely, t_K^{LI} and t_K^F (sometimes severely) under-reject for one-sided testing against $H_1 : \beta < 0$. Overall t_K^{LI} and t_K^F show similar performances, but t_K^{LI} is slightly better than t_K^F for testing against $H_1 : \beta < 0$.
- **iii):** Compared to t^{LI} and the (unadjusted) large-K t-ratios (t_K^{LI} and t_K^F), the proposed adjusted t-ratios ($t_K^{adj,LI}$ and $t_K^{adj,F}$) work well for all cases. Their rejection frequencies are overall close to the nominal level, and do not show undesirable asymmetries for one-sided testing as in the unadjusted test statistics t_K^{LI} and t_K^F. The performances of $t_K^{adj,LI}$ and $t_K^{adj,F}$ are similar.

[4]In our preliminary simulation, we also consider the t_3 and t_5 distributions as other examples of the elliptically contoured distribution. Although these distributions do not have finite 8th moments (i.e., the first condition in Assumption 2 (ii) is violated), we find that the simulation results are similar to the normal case.

Overall, the adjusted ratios, $t_K^{adj,LI}$ and $t_K^{adj,F}$, perform well for all cases. In particular, the adjustments improve the size distortions for one-sided testing because the null distributions of the adjusted t-ratios are less skewed and closer to the standard normal distribution, which agrees with our results of the asymptotic expansions in Section 2.

Power comparison. Next, we conduct power comparisons of the two-sided unadjusted and adjusted large-K t-tests. We generate 5,000 datasets from the DGP in (20) for various values of β and report power curves at 5% significance level. Figures 1-2 display the power curves. Among various cases tried in preliminary simulations, we present the cases of $\mu^2 = 30$ as typical examples. From these figures, we can see that: (i) due to the kinks of the power curves of the (unadjusted) large-K t-tests (t_K^{LI} and t_K^{F}) for small negative values of β, the adjusted large-K t-tests ($t_K^{adj,LI}$ and $t_K^{adj,F}$) have better local maximin power properties, and (ii) for negative (or positive) values of β, the adjusted large-K t-tests are more (or less) powerful than the unadjusted large-K t-tests.

Heteroskedastic errors. Finally, we study finite sample performances of the heteroskedastic versions of the large-K t-ratios discussed in Section 3.3. In particular, we consider heteroskedastic error terms with $u_i^H = (1 + 0.01 z_{2i}^{*2}) u_i$, where z_{2i}^* is the first element of z_{2i}. Tables 3 and 4 report the null rejection frequencies of the one-sided and two-sided tests at the nominal 5% significance level, respectively. We can see that the null distributions of the adjusted t-ratios reduce the skewness (asymmetries in the rejection frequencies for the positive and negative alternatives) and are closer to the standard normal distribution. Also our preliminary simulation suggests that the power curves are similar to the case of homoskedastic errors.

Appendix A. Proofs

A.1. **Proof of Theorem 2.** Let $\theta = (\beta', \gamma')'$, $\hat{\theta}_{LI} = (\hat{\beta}'_{LI}, \hat{\gamma}'_{LI})'$, $\hat{\theta}_F = (\hat{\beta}'_F, \hat{\gamma}'_F)'$,

$$\hat{e} = \begin{cases} \sqrt{n}(\hat{\theta}_{LI} - \theta) & \text{if} \quad \lambda = \hat{\lambda}, \\ \sqrt{n}(\hat{\theta}_F - \theta) & \text{if} \quad \lambda = \hat{\lambda} - a/(n-K), \end{cases}$$

and $X = [Y_2, Z_1]$. By the definitions of the LIML and Fuller estimators in (3), we have

$$X'(P - \lambda M)X\hat{e} = \sqrt{n}X'(P - \lambda M)(y, X)\begin{pmatrix} 1 \\ -\theta \end{pmatrix}. \tag{22}$$

Using the definition

$$D = (D_1, D_2) = \left(\begin{pmatrix} \pi_{11} \\ \pi_{21} \end{pmatrix}, \begin{pmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{pmatrix} \right),$$

it can be written as

$$\begin{aligned} & (Y, Z_1)'(P - \lambda M)(Y, Z_1) \\ = \ & \{ZD + (V, 0)\}'(P - \lambda M)\{ZD + (V, 0)\} \\ = \ & D'Z'ZD + D'Z'(V, 0) + (V, 0)'ZD + (V, 0)'(P_Z - \lambda M)(V, 0). \end{aligned} \tag{23}$$

Also we note that $\sqrt{K}(V'PV/K - \Omega) = O_p(1)$ and $\sqrt{n-K}(V'MV/(n-K) - \Omega) = O_p(1)$. By substituting (23) into (22) and putting

$$\begin{aligned} \hat{e} &= e^{(0)} + \frac{1}{\sqrt{n}}e^{(1)} + O_p(n^{-1}), \\ \hat{\lambda} &= \lambda^{(0)} + \frac{1}{\sqrt{n}}\lambda^{(1)} + \frac{1}{n}\lambda^{(2)} + O_p(n^{-3/2}), \end{aligned}$$

under Assumption 2 (i), we can determine successively $(e^{(0)}, e^{(1)})$ and $(\lambda^{(0)}, \lambda^{(1)}, \lambda^{(2)})$ as

$$\begin{aligned} \lambda^{(0)} &= c_*, \qquad \lambda^{(1)} = \frac{c_*}{\sigma^2}\left[\frac{\sqrt{n}}{K}u'Pu - \frac{\sqrt{n}}{n-K}u'Mu\right], \\ \lambda^{(2)} &= \frac{c_*}{\sigma^2}\left\{ -\frac{n}{K}e^{(0)\prime}Qe^{(0)} \right. \\ & \left. -\sqrt{\frac{n}{n-K}}\frac{1}{\sigma^2}\left[\frac{\sqrt{n}}{K}u'Pu - \frac{\sqrt{n}}{n-K}u'Mu\right]\left[\sqrt{n-K}(\frac{1}{n-K}u'Mu - \sigma^2)\right] \right\}, \\ e^{(0)} &= Q^{-1}\left[\frac{1}{\sqrt{n}}D_2'Z'u + \frac{\sqrt{c}}{\sqrt{K}}W_2'Pu - \frac{\sqrt{cc_*}}{\sqrt{n-K}}W_2'Mu\right], \\ e^{(1)} &= -Q^{-1}\left[\left\{\frac{1}{\sqrt{n}}D_2'Z'(V_2, 0) + \frac{\sqrt{c}}{\sqrt{K}}W_2'P(V_2, 0) - \frac{\sqrt{cc_*}}{\sqrt{n-K}}W_2'M(V_2, 0)\right\}e^{(0)} \right. \\ & + \frac{1}{\sqrt{n}}W_2'ZD_2e^{(0)} - (1-c)\lambda^{(1)}\left\{\begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2q_2'\sigma^2\right\}e^{(0)} \\ & \left. + \sqrt{1-c}\lambda^{(1)}\frac{1}{\sqrt{n-K}}W_2'Mu\right] + a(1+c_*)\sigma^2Q^{-1}q_2, \end{aligned} \tag{24}$$

where $W_2 = (V_2, 0) - uq_2'$, $q_2 = \frac{1}{\sigma^2}(\omega_{21}' - \beta'\Omega_{22}, 0')'$, and $c_* = c/(1-c)$. Each $\lambda^{(l)}$ is obtained by premultiplying $(1, -\beta', -\gamma')$ to (22). Each $e^{(l)}$ is obtained by using the last $G_1 + K_1$ rows of (22).

It should be noted that W_2 and u are uncorrelated when the rows of V are independently distributed. Moreover, we notice that $E[w_{2i}w_{2i}'w_{2i}u_i] = 0$ and $E[w_{2i}w_{2i}'w_{2i}u_i^3] = 0$ when W_2 and u follow some elliptically contoured distribution (by Assumption 2 (ii)). Using these facts, the Cornish-Fisher expansions of $\frac{1}{\sqrt{n}}D_2'Z'u$, $\frac{1}{\sqrt{K}}W_2'P_Z u$, and $\frac{1}{\sqrt{n-K}}W_2'Mu$ may be written as

$$\begin{aligned} \frac{1}{\sqrt{n}}D_2'Z'u &= \mathcal{X} + O_p(n^{-1}), \qquad \frac{1}{\sqrt{K}}W_2'Pu = \mathcal{Y} + O_p(n^{-1}), \\ \frac{1}{\sqrt{n-K}}W_2'Mu &= \mathcal{Z} + O_p(n^{-1}), \end{aligned}$$

where $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are some normally distributed random vectors. Thus, $\hat{e}$ can be written as

$$\hat{e} = \mathcal{E} + \frac{1}{\sqrt{n}}e^{(1)} + O_p(n^{-1}), \tag{25}$$

where $\mathcal{E} = Q^{-1}(\mathcal{X} + \sqrt{c}\mathcal{Y} + \sqrt{cc_*}\mathcal{Z})$.

We derive an asymptotic expansion of the distribution of $\hat{e}$ by inverting the characteristic function of $\hat{e}$ up to order $n^{-1/2}$:

$$E[\exp(\mathrm{i}s'\mathcal{E})] + \frac{1}{\sqrt{n}}E[\mathrm{i}s'E[e^{(1)}|\mathcal{E}]\exp(\mathrm{i}s'\mathcal{E})] + O(n^{-1}), \tag{26}$$

where s is a $(G_1+K_1)\times 1$ vector of real variables and $\mathrm{i} = \sqrt{-1}$. The conditional expectation of $e^{(1)}$ given the first order term $\mathcal{E}$ can be written as

$$E[e^{(1)}|\mathcal{E}] = -\{\mathcal{E}\mathcal{E}' - a(1+c_*)Q^{-1}\sigma^2\}q_2 + O_p(n^{-1/2}).$$

Therefore, the probability $P(\hat{e} \le \xi)$ is approximated to the order $n^{-1/2}$ by the Fourier inversion of the characteristic function (26). The inversion of the first term is $\Phi_\Psi(\xi)$. We also use the next Fourier inversion formula, which is a generalization of Fujikoshi *et al.* (1982): for $\mathcal{E} \sim N(\mu, \Sigma)$ and any polynomials $h(\cdot)$ and $g(\cdot)$,

$$\mathcal{F}^{-1}[h(-\mathrm{i}s)E[g(\mathcal{E})\exp(\mathrm{i}s'\mathcal{E})]]_{\mathcal{E}=\xi} = h\left(\frac{\partial}{\partial\xi}\right)g(\xi)\phi_{\mu,\Sigma}(\xi), \tag{27}$$

where $\phi_{\mu,\Sigma}(\xi)$ is the density function of $N(\mu,\Sigma)$ and $\partial/\partial\xi' = (\partial/\partial\xi_1, \cdots, \partial/\partial\xi_{G_1+K_1})$. The conclusion follows by applying this formula.

A.2. **Proof of Theorem 1.** To derive the asymptotic expansion of the null distribution of the large-K t-ratio, we first expand each term of $\hat{\Psi}$ in (8). Let E_1 and E_2 such that

$$\begin{aligned} \frac{1}{K}(V,0)'P(V,0) &= \begin{pmatrix} \Omega & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{K}}E_1, \\ \frac{1}{n-K}(V,0)'M(V,0) &= \begin{pmatrix} \Omega & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{n-K}}E_2. \end{aligned}$$

Using (23), and a $(1+G_2+K_1)\times(G_2+K_1)$ choice matrix $J_2=(0,I_{G_2+K_1})'$, we have

$$\begin{aligned}\hat{Q} &= Q+\frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{n}}D_2'Z'(V_2,0)+\frac{1}{\sqrt{n}}\begin{pmatrix}V_2'\\0'\end{pmatrix}ZD_2\right.\\ &\left.\quad+\sqrt{c}J_2'E_1J_2+\sqrt{cc_*}J_2'E_2J_2-\lambda^{(1)}\begin{pmatrix}\Omega_{22}&0\\0&0\end{pmatrix}\right]+O_p(n^{-1}),\end{aligned}$$

where the equality follows from Assumption 2 (iii). Thus $\hat{Q}^{-1}$ is expanded as

$$\hat{Q}^{-1}=Q^{-1}-\frac{1}{\sqrt{n}}Q^{-1}BQ^{-1}+O_p(n^{-1}),$$

where

$$B=\frac{1}{\sqrt{n}}D_2'Z'(V_2,0)+\frac{1}{\sqrt{n}}(V_2,0)'ZD_2+\sqrt{c}J_2'E_1J_2+\sqrt{cc_*}J_2'E_2J_2-\lambda^{(1)}\begin{pmatrix}\Omega_{22}&0\\0&0\end{pmatrix}.$$

Also, note that

$$\begin{aligned}&\frac{1}{n-K}\hat{b}'Y'MY\hat{b}\\ =&\left\{b-\frac{1}{\sqrt{n}}\begin{pmatrix}0\\e_\beta^{(0)}\end{pmatrix}+O_p(n^{-1})\right\}'\\ &\times\left\{\Omega+\frac{1}{\sqrt{n}}\left[\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}V'MV-\Omega\right)\right]\right\}\left\{b-\frac{1}{\sqrt{n}}\begin{pmatrix}0\\e_\beta^{(0)}\end{pmatrix}+O_p(n^{-1})\right\}\\ =&\ \sigma^2+\frac{1}{\sqrt{n}}\left[-2(0,e_\beta^{(0)\prime})\Omega b+\sqrt{\frac{n-K}{1-c}}b'\left(\frac{1}{n-K}V'MV-\Omega\right)b\right]+O_p(n^{-1}),\end{aligned}$$

where $b=(1,-\beta')'$ and $e_\beta^{(0)}$ is the first G_2 elements of $e^{(0)}$ in (24). Similarly,

$$\begin{aligned}&\frac{1}{(n-K)^2}Y'MY\hat{b}\hat{b}'Y'MY\\ =&\ \Omega bb'\Omega+\frac{1}{\sqrt{n}}\left[-\Omega b(0,e_\beta^{(0)\prime})\Omega+\sqrt{\frac{n-K}{1-c}}\Omega bb'\left(\frac{1}{n-K}V'MV-\Omega\right)\right.\\ &\left.-\Omega\begin{pmatrix}0\\e_\beta^{(0)}\end{pmatrix}b'\Omega+\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}V'MV-\Omega\right)bb'\Omega\right]+O_p(n^{-1}).\end{aligned}$$

Combining these terms, we have

$$\hat{\Psi}=\Psi+\frac{1}{\sqrt{n}}\Psi^{(1)}+O_p(n^{-1}),$$

where

$$\begin{aligned}
\Psi^{(1)} &= Q^{-1}\left[-2e^{(0)\prime}q_2\sigma^2+\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}u'Mu-\sigma^2\right)\right] \\
&\quad -Q^{-1}BQ^{-1}\sigma^2+(c_*+\kappa\eta)Q^{-1}AQ^{-1}-(c_*+\kappa\eta)Q^{-1}BQ^{-1}\left[\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}-q_2q_2'\sigma^4\right]Q^{-1} \\
&\quad -(c_*+\kappa\eta)Q^{-1}\left[\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}-q_2q_2'\sigma^4\right]Q^{-1}BQ^{-1}+\kappa^*\eta Q^{-1}\left[\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}-q_2q_2'\sigma^4\right]Q^{-1}, \\
\kappa^* &= \frac{1}{3\sigma^2}\left[-\frac{4}{n}\sum_{i=1}^{n}u_i^3w_{2i}'e^{(0)}-\frac{4}{n}\sum_{i=1}^{n}u_i^3z_i'D_2e^{(0)}\right. \\
&\quad \left.+\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}u_i^4-E[u_i^4]\right)-\frac{2E[u_i^4]}{\sigma^2}\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}u'Mu-\sigma^2\right)\right], \\
A &= -2\begin{pmatrix}\Omega_{22} & 0\\ 0 & 0\end{pmatrix}e^{(0)\prime}q_2\sigma^2+\sqrt{\frac{n-K}{1-c}}\begin{pmatrix}\Omega_{22} & 0\\ 0 & 0\end{pmatrix}\left(\frac{1}{n-K}u'Mu-\sigma^2\right) \\
&\quad +\sqrt{\frac{1}{1-c}}J_2'E_2J_2\sigma^2+q_2\sigma^2e^{(0)\prime}\begin{pmatrix}\Omega_{22} & 0\\ 0 & 0\end{pmatrix} \\
&\quad -q_2\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}u'M(V_2,0)-q_2'\sigma^2\right)\sigma^2+\begin{pmatrix}\Omega_{22} & 0\\ 0 & 0\end{pmatrix}e^{(0)}q_2'\sigma^2 \\
&\quad -\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}(V_2,0)'Mu-q_2\sigma^2\right)q_2'\sigma^2.
\end{aligned}$$

Under H_0, the large-K t-ratio (9) is approximated as

$$\begin{aligned}
t_K &= \frac{\hat{e}}{\sqrt{\Psi_j}}\sqrt{\frac{\Psi_j}{\hat{\Psi}_j}}=\frac{\iota'\left(e^{(0)}+\frac{1}{\sqrt{n}}e^{(1)}+O_p(n^{-1})\right)}{\sqrt{\Psi_j}}\left(1-\frac{1}{2\sqrt{n}}\frac{\Psi_j^{(1)}}{\Psi_j}+O_p(n^{-1})\right) \\
&= \mathcal{T}+\frac{1}{\sqrt{n}}t^{(1)}+O_p(n^{-1}), \qquad (28)
\end{aligned}$$

where

$$\mathcal{T}=\frac{\iota'e^{(0)}}{\sqrt{\Psi_j}},\qquad t^{(1)}=\frac{\iota'e^{(1)}}{\sqrt{\Psi_j}}-\frac{1}{2}\frac{\Psi_j^{(1)}}{\Psi_j}\mathcal{T}.$$

The first-order term $\mathcal{T}$ is asymptotically distributed as $N(0,1)$. We derive an asymptotic expansion of the distribution function of the large-K t-ratio by inverting the characteristic function up to $O(n^{-1/2})$:

$$E[\exp(\mathrm{i}s\mathcal{T})]+\frac{1}{\sqrt{n}}E[\mathrm{i}sE[t^{(1)}|\mathcal{T}]\exp(\mathrm{i}s\mathcal{T})]+O(n^{-1}).$$

By using Kunitomo and Matsushita (2009, Lemma A.3) and the fact that any odd moment of the elliptically contoured distribution is 0 (by Assumption 2 (ii)), the expectation of $t^{(1)}$ conditional on $\mathcal{T}$ is calculated as

$$\begin{aligned} E[t^{(1)}|\mathcal{T}] &= -\frac{1}{\sqrt{\Psi_j}}\left[(\iota'\Psi q_2)\mathcal{T}^2 - \frac{a}{1-c}\sigma^2\iota' Q^{-1}q_2\right] - \frac{1}{2}\left[-\frac{4}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\mathcal{T}^2\right] \\ &= \frac{1}{\sqrt{\Psi_j}}\left[\mathcal{T}^2\iota'\Psi + \frac{a}{1-c}\sigma^2\iota' Q^{-1}\right]q_2 + O_p(n^{-1/2}), \end{aligned} \tag{29}$$

where the first equality follows from Lemma 1. The probability $P\{t_K \le \xi\}$ is approximated to the order $n^{-1/2}$ using the same formula as (27).

The validity of the expansion is given by similar arguments to those in Kunitomo and Matsushita (2009) and in Fujikoshi *et al.* (1982). The random variables that appear in our analyses are $x_1 = \frac{1}{\sqrt{n}}D_2'Z'u$, $x_2 = \frac{1}{\sqrt{n}}D_2'Z'W_2$, $x_3 = \sqrt{K}(u'Pu/K - \sigma^2)$, $x_4 = \sqrt{n-K}(u'Mu/(n-K) - \sigma^2)$, $x_5 = \frac{1}{\sqrt{K}}W_2'Pu$, $x_6 = \frac{1}{\sqrt{n-K}}W_2'Mu$, $x_7 = \sqrt{K}(W_2'PW_2/K - C_2)$, and $x_8 = \sqrt{n-K}(W_2'MW_2/(n-K) - C_2)$, where $C_2 = E[w_{2i}w_{2i}']$. We use the space J_n where each element of x_i (for $i = 1, \ldots, 8$) is in the interval $(-2c\sqrt{\log n}, 2c\sqrt{\log n})$ and c is a standard deviation of each random variable. Then, if $E[||v_i||^8] < \infty$, we can take a positive constant d which satisfies

$$P\{||x_j|| > \sqrt{\Lambda_n \log n}\} \le \frac{d}{n(\log n)^2},$$

where Λ_n as the maximum of the characteristic roots of the covariance matrix of x_j $(j = 1, \ldots, 8)$ (Bhattacharya and Ghosh, 1978). Then, $P(J_n) = 1 - O(n^{-1})$, which can be proved in the same way as in Anderson (1974). We see that each element of $e^{(l)}$ and $t^{(l)}$ is a homogeneous polynomial of degree $l+1$ in the elements of x_j. The remainder terms of (25) and (28) are of the order $O(n^{-1})$ uniformly in J_n. Therefore, the analysis subsequent to (B.3) in Fujikoshi *et al.* (1982) is applicable.

Lemma 1. *Based on the setup and notation of the proof of Theorem 1, it holds* $E[\Psi_j^{(1)}|\mathcal{T}] = -4\sqrt{\Psi_j}(\iota'\Psi q_2) + O_p(n^{-1/2})$.

Proof of Lemma 1. Decompose

$$\begin{aligned} \Psi_j^{(1)} &= \iota'\{-2Q^{-1}Q^{(1)}\Psi + Q^{-1}\Sigma^{(1)}Q^{-1}\}\iota \\ &= -2\iota' Q^{-1}Q^{(1)}\Psi\iota + \iota' Q^{-1}\iota\left\{-2e^{(0)\prime}q_2\sigma^2 + \sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}u'Mu - \sigma^2\right)\right\} + \iota' Q^{-1}Q^{(1)}Q^{-1}\iota\sigma^2 \\ &\quad + (c_* + \kappa\eta)\iota' Q^{-1}AQ^{-1}\iota + \kappa^*\eta\iota' Q^{-1}\left\{\begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} - q_2 q_2'\sigma^4\right\}Q^{-1}\iota \\ &\equiv \Psi_{j1}^{(1)} + \Psi_{j2}^{(1)} + \Psi_{j3}^{(1)} + \Psi_{j4}^{(1)} + \Psi_{j5}^{(1)}. \end{aligned}$$

For $\Psi_{j5}^{(1)}$, we have $E[\Psi_{j5}^{(1)}|\mathcal{T}]=O_p(n^{-1/2})$. For $\Psi_{j1}^{(1)}$, we have

$$
\begin{aligned}
E[\Psi_{j1}^{(1)}|\mathcal{T}] &= -\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi\iota)(\iota'\Psi q_2)\mathcal{T}-\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi Q\Psi\iota)(\iota' Q^{-1}q_2)\mathcal{T}+O_p(n^{-1/2}) \\
&= -\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi\iota)(\iota'\Psi q_2)\mathcal{T} \\
&\quad -\frac{2}{\sqrt{\Psi_j}}\left(\iota' Q^{-1}\left[Q\sigma^2+(c_*+\kappa\eta)\left\{\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}-q_2q_2'\sigma^4\right\}\right]\Psi\iota\right)(\iota' Q^{-1}q_2)\mathcal{T}+O_p(n^{-1/2}) \\
&= -\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi\iota)(\iota'\Psi q_2)\mathcal{T}-\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi\iota)(\iota' Q^{-1}\sigma^2 q_2)\mathcal{T} \\
&\quad -\frac{2}{\sqrt{\Psi_j}}(c_*+\kappa\eta)\left\{\iota' Q^{-1}\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}\Psi\iota\right\}(\iota' Q^{-1}q_2)\mathcal{T} \\
&\quad +\frac{2}{\sqrt{\Psi_j}}(c_*+\kappa\eta)(\iota' Q^{-1}q_2q_2'\sigma^4 Q^{-1}\iota)(\iota'\Psi q_2)\mathcal{T}+O_p(n^{-1/2}),
\end{aligned}
\tag{30}
$$

For $\Psi_{j2}^{(1)}$, we have

$$
E[\Psi_{j2}^{(1)}|\mathcal{T}]=-\frac{2}{\sqrt{\Psi_j}}(\iota' Q^{-1}\sigma^2\iota)(\iota' E[e^{(0)}e^{(0)\prime}]q_2)\mathcal{T}+O_p(n^{-1/2})=-\frac{2}{\sqrt{\Psi_j}}(\iota' Q^{-1}\sigma^2\iota)(\iota'\Psi q_2)\mathcal{T}+O_p(n^{-1/2}).
\tag{31}
$$

For $\Psi_{j3}^{(1)}$, we have

$$
E[\Psi_{j3}^{(1)}|\mathcal{T}]=E[\iota' Q^{-1}Q^{(1)}Q^{-1}\iota\sigma^2|\mathcal{T}]=\frac{2}{\sqrt{\Psi_j}}(\iota'\Psi\iota)(\iota' Q^{-1}\sigma^2 q_2)\mathcal{T}+O_p(n^{-1/2}).
\tag{32}
$$

For $\Psi_{j4}^{(1)}$, note that

$$
\begin{aligned}
\iota' Q^{-1}AQ^{-1}\iota &= \iota' Q^{-1}\left[-2\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}e^{(0)\prime}q_2+\sqrt{\frac{n-K}{1-c}}\begin{pmatrix}\Omega_{22} & 0\\ 0 & 0\end{pmatrix}\left(\frac{1}{n-K}u'Mu-\sigma^2\right)\right. \\
&\quad +\sqrt{\frac{1}{1-c}}J_2'E_2J_2\sigma^2+q_2e^{(0)\prime}\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix} \\
&\quad -q_2\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}u'M(V_2,0)-q_2'\sigma^2\right)\sigma^2+\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}e^{(0)}q_2' \\
&\quad \left.-\sqrt{\frac{n-K}{1-c}}\left(\frac{1}{n-K}(V_2,0)'Mu-q_2\sigma^2\right)q_2'\sigma^2\right]Q^{-1}\iota \\
&\equiv A_{j1}+\cdots+A_{j7}.
\end{aligned}
$$

Observe that $E[A_{j2}+A_{j3}+A_{j5}+A_{j7}|\mathcal{T}]=O_p(n^{-1/2})$, and

$$
\begin{aligned}
E[A_{j1}|\mathcal{T}] &= -\frac{2}{\sqrt{\Psi_j}}\left(\iota' Q^{-1}\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}Q^{-1}\iota\right)(\iota'\Psi q_2)\mathcal{T}+O_p(n^{-1/2}), \\
E[A_{j4}+A_{j6}|\mathcal{T}] &= \frac{2}{\sqrt{\Psi_j}}\left(\iota' Q^{-1}\begin{pmatrix}\sigma^2\Omega_{22} & 0\\ 0 & 0\end{pmatrix}\Psi\iota\right)(\iota' Q^{-1}q_2)\mathcal{T}+O_p(n^{-1/2}).
\end{aligned}
$$

Combining these results,

$$
\begin{aligned}
E[\Psi_{j4}^{(1)}|\mathcal{T}] &= -\frac{2}{\sqrt{\Psi_j}}(c_* + \kappa\eta)\left(\iota' Q^{-1}\begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}\iota\right)(\iota'\Psi q_2)\mathcal{T} \\
&\quad + \frac{2}{\sqrt{\Psi_j}}(c_* + \kappa\eta)\left(\iota' Q^{-1}\begin{pmatrix} \sigma^2\Omega_{22} & 0 \\ 0 & 0 \end{pmatrix} \Psi\iota\right)(\iota' Q^{-1} q_2)\mathcal{T} + O_p(n^{-1/2}). \quad (33)
\end{aligned}
$$

Combining (30)-(33), the conclusion follows. Note that the sum of the last term of (30), (31) and the first term of (33) equals $-2\sqrt{\Psi}_j(\iota'\Psi q_2)$.

A.3. **Proof of Theorem 3.** Under H_{1n}, the large-K t-ratio can be written as

$$
\begin{aligned}
t_K &= \frac{1}{\sqrt{\hat{\Psi}_j}}\iota'\sqrt{n}\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \frac{1}{\sqrt{\hat{\Psi}_j}}\iota'\sqrt{n}\begin{pmatrix} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{pmatrix} + \frac{\iota'\zeta}{\sqrt{\hat{\Psi}_j}} \\
&= \left(\mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}\right) + \frac{1}{\sqrt{n}}\left(t^{(1)} - \frac{\Psi_j^{(1)}\iota'\zeta}{2\Psi_j\sqrt{\Psi_j}}\right) + O_p(n^{-1}), \quad (34)
\end{aligned}
$$

where $\mathcal{T}$ in the first-order term is distributed as $N(0,1)$. Since the expectation of $\Psi_j^{(1)}$ conditional on $\mathcal{T}$ is calculated as $E[\Psi_j^{(1)}|\mathcal{T}] = -4\sqrt{\Psi_j}(\iota'\Psi q_2)\mathcal{T} + O_p(n^{-1/2})$ by Lemma 1, we have

$$
\begin{aligned}
& E\left[t^{(1)} - \frac{\Psi_j^{(1)}\iota'\zeta}{2\Psi_j\sqrt{\Psi_j}}\,\middle|\, \mathcal{T}_* = \mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}\right] \\
&= \frac{1}{\sqrt{\Psi_j}}\left[\mathcal{T}_*^2\iota'\Psi + \frac{a}{1-c}\sigma^2\iota' Q^{-1}\right]q_2 - \frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2 + O_p(n^{-1/2}). \quad (35)
\end{aligned}
$$

Then the probability $P\{t_K \le \tau\}$ is approximated to the order $O(n^{-1/2})$ using the inversion formula (27):

$$
\begin{aligned}
& P\{t_K < \tau\} \\
&= \Phi_\zeta(\tau) - \frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{\Psi_j}}\left\{\tau^2(\iota'\Psi) + \frac{a}{1-c}\sigma^2 Q^{-1}\right\}q_2 - \frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right]\phi_\zeta(\tau) + O(n^{-1}).
\end{aligned}
$$

The result for the two-sided test is obtained by $P\{|t_K| \le \tau\} = P\{t_K \le \tau\} - P\{t_K \le -\tau\}$.

A.4. **Proof of Theorems 4 and 5.** We only present the proof for Theorem 5 since the proof of Theorem 4 for the null distribution follows directly by setting $\zeta = 0$.

By using (34), the adjusted t-ratio under H_{1n} can be written as

$$
\begin{aligned}
t_K^{adj} &= t_K - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\hat{\Psi}_j}}\left\{(\iota'\hat{\Psi}\hat{q}_2)t_{K0}^2 + \frac{an}{n-K}\hat{\sigma}^2\iota'\hat{Q}^{-1}\hat{q}_2\right\} + O_p(n^{-1}) \\
&= \left(\mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}\right) + \frac{1}{\sqrt{n}}\left(t^{(1)} - \frac{\Psi_j^{(1)}\iota'\zeta}{2\Psi_j\sqrt{\Psi_j}}\right) \\
&\quad - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2 - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{\Psi_j}}\frac{a}{1-c}\sigma^2\iota' Q^{-1} q_2 + O_p(n^{-1}).
\end{aligned}
$$

Thus, using (35), the expectation of the $O(n^{-1/2})$ term conditional on $\mathcal{T}_* = \mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}$ is calculated as

$$
\begin{aligned}
& E\left[\left(t^{(1)} - \frac{\Psi_j^{(1)}\iota'\zeta}{2\Psi_j\sqrt{\Psi_j}}\right) - \frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\mathcal{T} + \frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2 - \frac{1}{\sqrt{\Psi_j}}\frac{a}{1-c}\sigma^2\iota'Q^{-1}q_2 \middle| \mathcal{T}^*\right] \\
= \; & -\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2 + O_p(n^{-1/2}). \qquad (36)
\end{aligned}
$$

The probability $P\{t_K^{adj} \leq \tau\}$ is approximated to the order $n^{-1/2}$ using the inversion formula (27):

$$
P\{t_K^{adj} \leq \tau\} = \Phi_\zeta(\tau) - \frac{1}{\sqrt{n}}\left[-\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\left(\frac{\iota'\zeta}{\sqrt{\Psi_j}}\right)^2\right]\phi_\zeta(\tau) + O(n^{-1}).
$$

The result for the two-sided test is obtained by $P\{|t_K^{adj}| \leq \tau\} = P\{t_K^{adj} \leq \tau\} - P\{t_K^{adj} \leq -\tau\}$.

A.5. **Proof of Proposition 1.** By Theorem 3 and Theorem 5, we obtain

$$
\begin{aligned}
& \lim_{n\to\infty}\sqrt{n}\left[\min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} P\{|t_K^{adj}| \geq \tau\} - \min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} P\{|t_K^F| \geq \tau\}\right] \\
= \; & \min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} -\frac{1}{\sqrt{\Psi_j}}(\iota'\Psi q_2)\frac{\delta}{\Psi_j}\{\phi_\zeta(\tau) - \phi_\zeta(-\tau)\} \\
& - \min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta} -\frac{1}{\sqrt{\Psi_j}}\left\{(\iota'\Psi q_2)\left(\frac{\delta}{\Psi_j} - \tau^2\right) - \frac{a\sigma^2}{1-c}\iota'Q^{-1}q_2\right\}\{\phi_\zeta(\tau) - \phi_\zeta(-\tau)\}.
\end{aligned}
$$

Since $\min_{\zeta\in\mathbb{R}^{G_2+K_1}:(\iota'\zeta)^2=\delta}\{\phi_\zeta(\tau) - \phi_\zeta(-\tau)\}$ is negative, the right hand side becomes positive as far as $|\iota'\Psi q_2|\frac{\delta}{\Psi_j} < \left|(\iota'\Psi q_2)\left(\frac{\delta}{\Psi_j} - \tau^2\right) - \frac{a\sigma^2}{1-c}\iota'Q^{-1}q_2\right|$ or equivalently $\delta \in \left(0, \left|\frac{\tau^2\Psi_j}{2} + \frac{\frac{a\sigma^2}{1-c}(\iota'Q^{-1}q_2)\Psi_j}{2(\iota'\Psi q_2)}\right|\right)$ by using the fact that if $|a| < |b|/2$, then $|a-b| - |a| > (|b| - |a|) - |b|/2 > |b|/2 - |a| > 0$. Thus, the conclusion follows by setting $\Delta = \left|\frac{\tau^2\Psi_j}{2} + \frac{\frac{a\sigma^2}{1-c}(\iota'Q^{-1}q_2)\Psi_j}{2(\iota'\Psi q_2)}\right|$.

A.6. **Proof of Proposition 2.** Let $\hat{e}_H = \sqrt{n}(\hat{\theta}_{HLIM} - \theta)$ and $X = (Y_2, Z_1)$. By the definition of the HLIM estimator in (15), we have

$$
(y_1, X)'(P^* - \hat{\alpha}I)X\hat{e}_H = \sqrt{n}(y_1, X)'(P^* - \hat{\alpha}I)(y_1, X)\begin{pmatrix} 1 \\ -\theta \end{pmatrix}. \qquad (37)
$$

By the definition $D = (D_1, D_2) = \left(\begin{pmatrix} \pi_{11} \\ \pi_{21} \end{pmatrix}, \begin{pmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{pmatrix}\right)$, it can be written as

$$
\begin{aligned}
& (Y, Z_1)'(P^* - \hat{\alpha}I)(Y, Z_1) \\
= \; & \{ZD + (V, 0)\}'(P^* - \hat{\alpha}I)\{ZD + (V, 0)\} \\
= \; & D'\sum_{i=1}^n z_i z_i'(1 - P_{ii})D + D'\sum_{i=1}^n z_i(1 - P_{ii})(v_i', 0) + \sum_{i=1}^n (v_i', 0)'(1 - P_{ii})z_i D + (V, 0)'P^*(V, 0) \\
& -\hat{\alpha}\{ZD + (V, 0)\}'\{ZD + (V, 0)\}. \qquad (38)
\end{aligned}
$$

Also we note that $V'P^*V/\sqrt{n} = O_p(1)$ and $\sqrt{n}(V'V/n - \bar{\Omega}) = O_p(1)$, where $\bar{\Omega} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E[v_i v_i']$. By substituting (38) into (37) and putting

$$
\begin{aligned}
\hat{e}_H &= e_H^{(0)} + \frac{1}{\sqrt{n}} e_H^{(1)} + O_p(n^{-1}), \\
\hat{\alpha} &= \alpha^{(0)} + \frac{1}{\sqrt{n}}\alpha^{(1)} + \frac{1}{n}\alpha^{(2)} + O_p(n^{-3/2}),
\end{aligned}
\tag{39}
$$

under Assumption 2 (i), we can determine successively $(e_H^{(0)}, e_H^{(1)})$ and $(\alpha^{(0)}, \alpha^{(1)}, \alpha^{(2)})$. Let $w_{2i} = (v_{2i}', 0)' - u_i \bar{q}_2$, $\bar{q}_2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E[(v_{2i}', 0)' u_i]/\bar{\sigma}^2$, $\bar{\sigma}^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E[u_i^2]$, $\bar{\Omega}_{22} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E[v_{2i}v_{2i}']$, and $\bar{Q} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n D_2' z_i z_i' D_2$. For $\hat{\alpha}$, by premultiplying $(1, -\beta', -\gamma')$ to (37), we have

$$
\begin{aligned}
\hat{\alpha} &= \frac{1}{\sqrt{n}} \frac{\frac{1}{\sqrt{n}} u'P^*u - \frac{1}{\sqrt{n}}\left(\frac{1}{\sqrt{n}} u'P^*X\right)\hat{e}_H}{\bar{\sigma}^2 + \left(\frac{1}{n}u'u - \bar{\sigma}^2\right) - \frac{\bar{\sigma}^2}{\sqrt{n}}\bar{q}_2'\hat{e}_H - \frac{1}{\sqrt{n}}\left(\frac{1}{n}u'X - \bar{\sigma}^2\bar{q}_2'\right)\hat{e}_H} \\
&= \frac{1}{\sqrt{n}}\left[\frac{1}{\bar{\sigma}^2}\frac{1}{\sqrt{n}} u'P^*u\right] \\
&\quad + \frac{1}{n}\left[-\frac{1}{\bar{\sigma}^2}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n u_i(1-P_{ii})z_i'D_2 + \frac{1}{\sqrt{n}}u'P^*V - \frac{1}{\sqrt{n}}u'P^*u\bar{q}_2'\right)e_H^{(0)}\right] \\
&\quad - \frac{1}{n}\left[\frac{1}{\bar{\sigma}^2}\left(\frac{1}{\sqrt{n}}u'P^*u\right)\frac{\sqrt{n}}{\bar{\sigma}^2}\left(\frac{1}{n}u'u - \bar{\sigma}^2\right)\right] + O_p(n^{-3/2}) \\
&= \frac{1}{\sqrt{n}}\left[\frac{1}{\bar{\sigma}^2}\frac{1}{\sqrt{n}}u'P^*u\right] + \frac{1}{n}\left[-\frac{1}{\bar{\sigma}^2}e_H^{(0)\prime}Q_H e_H^{(0)} - \frac{1}{\bar{\sigma}^2}\left(\frac{1}{\sqrt{n}}u'P^*u\right)\frac{\sqrt{n}}{\bar{\sigma}^2}\left(\frac{1}{n}u'u - \bar{\sigma}^2\right)\right] + O_p(n^{-3/2}),
\end{aligned}
$$

where the second equality follows from a Taylor expansion and direct calculation, and the third equality is due to the definition

$$
e_H^{(0)} = Q_H^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2' z_i(1-P_{ii})u_i + \frac{1}{\sqrt{n}}\sum_{i\neq j}^n P_{ij}w_{2i}u_j\right\}. \tag{40}
$$

For $\hat{e}_H$, by using the last G_2+K_1 rows of (37), we have

$$
\begin{aligned}
\hat{e}_H &= \left\{\frac{1}{n}X'(P^*-\hat{\alpha}I)X\right\}^{-1}\frac{1}{\sqrt{n}}X'(P^*-\hat{\alpha}I)u \\
&= Q_H^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2'z_i(1-P_{ii})u_i+\frac{1}{\sqrt{n}}\sum_{i\neq j}^n P_{ij}w_{2i}u_j\right\}-\alpha^{(1)}Q_H^{-1}\frac{1}{n}\sum_{i=1}^n w_{2i}u_i+\frac{1}{\sqrt{n}}e_H^{(0)\prime}Q_He_H^{(0)}Q_H\bar{q}_2 \\
&\quad -\frac{1}{\sqrt{n}}Q_H^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2'z_i(1-P_{ii})(v_{2i}',0)+\frac{1}{\sqrt{n}}\sum_{i=1}^n(v_{2i}',0)'(1-P_{ii})z_i'D_2\right. \\
&\qquad \left.+\frac{1}{\sqrt{n}}(V_2,0)'P^*(V_2,0)-\alpha^{(1)}\left\{\bar{Q}+\begin{pmatrix}\bar{\Omega}_{22} & 0\\ 0 & 0\end{pmatrix}\right\}\right]e_H^{(0)}+O_p(n^{-1}) \\
&= Q_H^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2'z_i(1-P_{ii})u_i+\frac{1}{\sqrt{n}}\sum_{i\neq j}^n P_{ij}w_{2i}u_j\right\}-\alpha^{(1)}Q_H^{-1}\frac{1}{n}\sum_{i=1}^n w_{2i}u_i \\
&\quad -\frac{1}{\sqrt{n}}Q_H^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2'z_i(1-P_{ii})(v_{2i}',0)+\frac{1}{\sqrt{n}}\sum_{i\neq j}P_{ij}w_{2i}(v_{2j}',0)\right. \\
&\qquad \left.+\frac{1}{\sqrt{n}}\sum_{i=1}^n w_{2i}'(1-P_{ii})z_i'D_2-\alpha^{(1)}\left\{\bar{Q}+\begin{pmatrix}\bar{\Omega}_{22} & 0\\ 0 & 0\end{pmatrix}-\bar{q}_2\bar{q}_2'\right\}\right]e_H^{(0)}+O_p(n^{-1}),
\end{aligned}
$$

where the second equality follows from the assumption on Q_H and (39) with the expression

$$
\alpha^{(1)}=\frac{1}{\bar{\sigma}^2}\frac{1}{\sqrt{n}}u'P^*u,\qquad \alpha^{(2)}=-\frac{1}{\bar{\sigma}^2}e^{(0)\prime}Q_ne^{(0)}-\frac{1}{\bar{\sigma}^2}\left(\frac{1}{\sqrt{n}}u'P^*u\right)\frac{\sqrt{n}}{\bar{\sigma}^2}\left(\frac{1}{n}u'u-\bar{\sigma}^2\right) \tag{41}
$$

and the third equality follows from the definition of $e_H^{(0)}$ in (40).

Combining these results, we obtain the expansion in (39) with $e_H^{(0)}$ in (40), $\alpha^{(1)}$ and $\alpha^{(2)}$ in (41), $\alpha^{(0)}=0$,

$$
\begin{aligned}
e_H^{(1)} &= -\alpha^{(1)}Q^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n w_{2i}u_i-\frac{1}{\sqrt{n}}Q^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n D_2'z_i(1-P_{ii})(v_{2i}',0)+\frac{1}{\sqrt{n}}\sum_{i\neq j}P_{ij}w_{2i}(v_{2j}',0)\right. \\
&\qquad \left.+\frac{1}{\sqrt{n}}\sum_{i=1}^n w_{2i}'(1-P_{ii})z_i'D_2-\alpha^{(1)}\left\{\bar{Q}+\begin{pmatrix}\bar{\Omega}_{22} & 0\\ 0 & 0\end{pmatrix}-\bar{q}_2\bar{q}_2'\right\}\right]e_H^{(0)}
\end{aligned} \tag{42}
$$

Therefore, the conclusion follows from

$$
\begin{aligned}
t_K^H &= \frac{\hat{e}_H}{\sqrt{\Psi_{H,j}}}\sqrt{\frac{\Psi_{H,j}}{\hat{\Psi}_{H,j}}}=\frac{\iota'\left(e_H^{(0)}+\frac{1}{\sqrt{n}}e_H^{(1)}+O_p(n^{-1})\right)}{\sqrt{\Psi_{H,j}}}\left(1-\frac{1}{2\sqrt{n}}\frac{\Psi_{H,j}^{(1)}}{\Psi_{H,j}}+O_p(n^{-1})\right) \\
&= \mathcal{T}_H+\frac{1}{\sqrt{n}}t_H^{(1)}+O_p(n^{-1}),
\end{aligned} \tag{43}
$$

where

$$
\mathcal{T}_H=\frac{\iota'e_H^{(0)}}{\sqrt{\Psi_{H,j}}},\qquad t_H^{(1)}=\frac{\iota'e_H^{(1)}}{\sqrt{\Psi_{H,j}}}-\frac{1}{2}\frac{\Psi_{H,j}^{(1)}}{\Psi_{H,j}}\mathcal{T}_H.
$$

A.7. **Proof of Proposition 3.** Lengthy calculations yield that

$$\begin{aligned}
E[\iota' e_H^{(1)}|\mathcal{T}_H] &= -(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H^2 + O_p(n^{-1/2}), \\
E[\Psi_{H,j}^{(1)}|\mathcal{T}_H] &= -2E[\iota' Q_H^{-1}Q_H^{(1)}\Psi_H\iota|\mathcal{T}_H] + E[\iota' Q_H^{-1}\Sigma_H^{(1)}Q_H^{-1}\iota|\mathcal{T}_H], \\
-2E[\iota' Q_H^{-1}Q_H^{(1)}\Psi_H\iota|\mathcal{T}_H] &= -2\sqrt{\Psi_{H,j}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H - \frac{2}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H Q_H\Psi_H\iota)(\iota' Q_H^{-1}\bar{q}_2)\mathcal{T}_H + O_p(n^{-1/2}), \\
E[\iota' Q_H^{-1}\Sigma_H^{(1)}Q_H^{-1}\iota|\mathcal{T}_H] &= -2\sqrt{\Psi_{H,j}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H \\
&\quad + \frac{2}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H Q_H\Psi_H\iota)(\iota' Q_H^{-1}\bar{q}_2)\mathcal{T}_H + O_p(n^{-1/2}). \qquad (44)
\end{aligned}$$

Since proofs are similar, we only present the proof of (44) below. Based on these conditional expectations, we obtain

$$\begin{aligned}
E[t_H^{(1)}|\mathcal{T}_H] &= E\left[\left.\frac{\iota' e_H^{(1)}}{\sqrt{\Psi_{H,j}}} - \frac{1}{2}\frac{\Psi_{H,j}^{(1)}}{\Psi_{H,j}}\mathcal{T}_H\right|\mathcal{T}_H\right] \\
&= -\frac{1}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H^2 - \frac{1}{2}\left[-\frac{4}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H^2\right] + O_p(n^{-1/2}) \\
&= \frac{1}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H^2 + +O_p(n^{-1/2}).
\end{aligned}$$

Therefore, we obtain an analogous expression for the conditional expectation in (29), and the conclusion follows by the same argument in the proof of Theorem 2 using the inversion formula in (27).

Proof of (44). Decompose

$$\begin{aligned}
\Sigma_H^{(1)} &= \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^n\sum_{i\neq k}^n\sum_{j\neq k}^n \hat{X}_i P_{ik}\hat{u}_k^2 P_{kj}\hat{X}_j' + \frac{1}{n}\sum_{i=1}^n\sum_{j\neq i}^n \hat{X}_i\hat{X}_j'\hat{u}_i\hat{u}_j P_{ij}^2 - \Sigma_H\right) \\
&\equiv S_0 + S_1 + S_2,
\end{aligned}$$

where

$$
\begin{aligned}
S_0 &= \sqrt{n}\left\{\frac{1}{n}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}(D_2' z_i + w_{2i})P_{ik}u_k^2 P_{kj}(D_2' z_j + w_{2j})' \right.\\
&\qquad \left. + \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(D_2' z_i + w_{2i})(D_2' z_j + w_{2j})' u_i u_j P_{ij}^2 - \Sigma_H\right\},\\
S_1 &= \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}(D_2' z_i + w_{2i})P_{ik}(\hat{u}_k^2 - u_k^2)P_{kj}(D_2' z_j + w_{2j})'\\
&\quad - \frac{2}{\sqrt{n}}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}\bar{q}_2(\hat{u}_i - u_i)P_{ik}u_k^2 P_{kj}(D_2' z_j + w_{2j})'\\
&\quad - \frac{2}{\sqrt{n}}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}(\hat{q}_2 - \bar{q}_2)u_i P_{ik}u_k^2 P_{kj}(D_2' z_j + w_{2j})'\\
&\equiv S_{11} + S_{12} + S_{13},\\
S_2 &= \frac{2}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(D_2' z_i + w_{2i})(D_2' z_j + w_{2j})'(\hat{u}_i - u_i)u_j P_{ij}^2 - \frac{2}{\sqrt{n}}\sum_{i\neq j}^{n}\bar{q}_2(\hat{u}_i - u_i)(D_2' z_j + w_{2j})' u_i u_j P_{ij}^2\\
&\quad - \frac{2}{\sqrt{n}}\sum_{i\neq j}^{n}(\hat{q}_2 - \bar{q}_2)u_i(D_2' z_j + w_{2j})' u_i u_j P_{ij}^2\\
&\equiv S_{21} + S_{22} + S_{23}.
\end{aligned}
$$

Then we can obtain

$$
\begin{aligned}
&\sqrt{n}E[\iota' Q_H^{-1}(S_{11} + S_{21})Q_H^{-1}\iota|\mathcal{T}_H]\\
= {}& E\left[\iota' Q_H^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}(D_2' z_i + w_{2i})P_{ik}\left(-\frac{2}{\sqrt{n}}u_k X_k' e_H^{(0)}\right)P_{kj}(D_2' z_j + w_{2j})'\right\}Q_H^{-1}\iota\right|\mathcal{T}_H\Bigg]\\
&+E\left[\iota' Q_H^{-1}\left\{\frac{2}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(D_2' z_i + w_{2i})(D_2' z_j + w_{2j})'\left(-\frac{1}{\sqrt{n}}X_i' e_H^{(0)}\right)u_j P_{ij}^2\right\}Q_H^{-1}\iota\right|\mathcal{T}_H\Bigg] + O_p(n^{-1/2})\\
= {}& E\left[\iota' Q_H^{-1}\left\{-\frac{2}{n}\sum_{i,j,k,i\neq j,j\neq k}^{n}\sigma_k^2 D_2' z_i P_{ik}P_{kj}z_j' D_2\right\}Q_H^{-1}\iota(\bar{q}_2' e_H^{(0)})\right|\mathcal{T}_H\Bigg]\\
&+E\left[\iota' Q_H^{-1}\left\{-\frac{2}{n}\sum_{i\neq j}^{n}P_{ij}^2\{E[w_{2i}w_{2i}']\sigma_j^2 + E[w_{2i}u_i]E[w_{2j}u_j]\right\}Q_H^{-1}\iota(\bar{q}_2' e_H^{(0)})\right|\mathcal{T}_H\Bigg] + O_p(n^{-1/2})\\
= {}& -2\sqrt{\Psi_{H,j}}(\iota'\Psi_H\bar{q}_2)\mathcal{T}_H + O_p(n^{-1/2}),
\end{aligned}
$$

where the first equality follows by $\hat{u}_k = u_k - \frac{1}{\sqrt{n}}X_k' e_H^{(0)} + O_p(n^{-1})$ and Chebyshev's inequality, the second equality follows from $X_i = D_2' z_i + w_{2i} + u_i\bar{q}_2$, Assumption 2H, and the assumption $\frac{1}{n}\sum_{i=1}^{n}(\Omega_i - \bar{\Omega})\otimes(\Omega_i - \bar{\Omega}) = o_p(1)$, and the third equality follows from the definition of Ψ_H. By similar arguments, we can obtain

$$
\begin{aligned}
&\sqrt{n}E[\iota' Q_H^{-1}(S_{12}+S_{22})Q^{-1}\iota|\mathcal{T}_H] \\
= \;& E\left[\iota' Q_H^{-1}\left\{-\frac{2}{\sqrt{n}}\sum_{k=1}^{n}\sum_{i\neq k}^{n}\sum_{j\neq k}^{n}\bar{q}_2\left(-\frac{1}{\sqrt{n}}X_i' e_H^{(0)}\right)P_{ik}u_k^2 P_{kj}(D_2' z_j + w_{2j})'\right\}Q_H^{-1}\iota\right|\mathcal{T}_H\Bigg] \\
&+E\left[\iota' Q_H^{-1}\left\{-\frac{2}{\sqrt{n}}\sum_{i\neq j}^{n}\bar{q}_2\left(-\frac{1}{\sqrt{n}}X_i' e_H^{(0)}\right)(D_2' z_j + w_{2j})' u_i u_j P_{ij}^2\right\}Q_H^{-1}\iota\right|\mathcal{T}_H\Bigg] \\
= \;& \frac{2}{\sqrt{\Psi_{H,j}}}(\iota'\Psi_H Q_H \Psi_H \iota)(\iota' Q_H^{-1}\bar{q}_2)\mathcal{T}_H + O_p(n^{-1/2}),
\end{aligned}
$$

and $\sqrt{n}E[\iota' Q_H^{-1}(S_{13}+S_{23})Q_H^{-1}\iota|\mathcal{T}_H] = O_p(n^{-1/2})$. Combining these results, we obtain (44).

Appendix B. Tables and figures

			t^{LI}		t_K^{LI}		t_K^F		$t_K^{adj,LI}$		$t_K^{adj,F}$	
ρ	μ^2	K	$\beta<0$	$\beta>0$	$\beta<0$	$\beta>0$	$\beta<0$	$\beta>0$	$\beta<0$	$\beta>0$	$\beta<0$	$\beta>0$
0.4	60	5	0.033	0.063	0.029	0.058	0.023	0.060	0.052	0.050	0.051	0.050
	60	10	0.039	0.072	0.028	0.060	0.022	0.062	0.054	0.053	0.053	0.052
	60	20	0.049	0.088	0.026	0.061	0.020	0.063	0.055	0.055	0.053	0.054
	30	5	0.024	0.073	0.016	0.061	0.010	0.063	0.057	0.051	0.052	0.055
	30	10	0.029	0.088	0.014	0.063	0.008	0.064	0.055	0.058	0.051	0.058
	30	20	0.043	0.114	0.011	0.064	0.008	0.066	0.056	0.063	0.051	0.063
0.8	60	5	0.018	0.077	0.016	0.075	0.009	0.085	0.055	0.055	0.051	0.055
	60	10	0.018	0.078	0.014	0.073	0.008	0.083	0.056	0.053	0.053	0.052
	60	20	0.022	0.087	0.014	0.075	0.008	0.085	0.055	0.053	0.052	0.052
	30	5	0.003	0.089	0.002	0.083	0.001	0.097	0.061	0.058	0.051	0.057
	30	10	0.003	0.096	0.002	0.085	0.003	0.098	0.061	0.058	0.050	0.057
	30	20	0.004	0.110	0.001	0.086	0.000	0.098	0.058	0.058	0.047	0.057

Table 1. Null rejection frequencies of one-sided large-K t-tests at 5% significance level (homoskedastic case)

ρ	μ^2	K	t^{LI}	t_K^{LI}	t_K^F	$t_K^{adj,LI}$	$t_K^{adj,F}$
0.4	60	5	0.048	0.042	0.043	0.055	0.054
	60	10	0.056	0.042	0.043	0.057	0.056
	60	20	0.073	0.041	0.042	0.058	0.057
	30	5	0.047	0.036	0.037	0.062	0.058
	30	10	0.061	0.037	0.037	0.064	0.060
	30	20	0.089	0.038	0.038	0.068	0.063
0.8	60	5	0.052	0.049	0.056	0.061	0.058
	60	10	0.052	0.047	0.054	0.060	0.056
	60	20	0.059	0.048	0.054	0.058	0.056
	30	5	0.058	0.054	0.063	0.069	0.059
	30	10	0.064	0.055	0.064	0.068	0.059
	30	20	0.075	0.056	0.064	0.065	0.055

Table 2. Null rejection frequencies of two-sided large-K t-tests at 5% significance level (homoskedastic case)

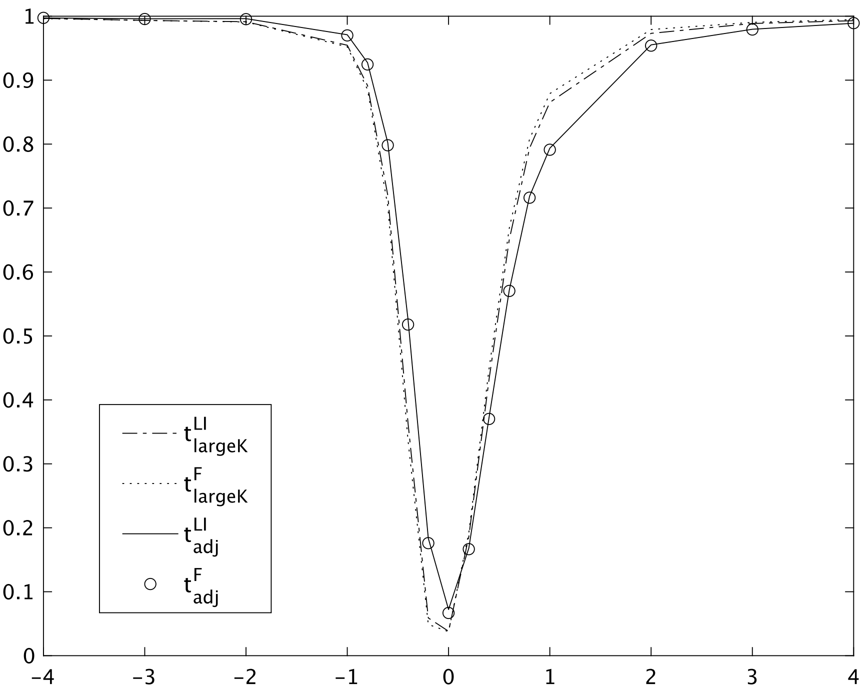

FIGURE 1. Power curves of two-sided tests: $n = 200$ (sample size), $K = 20$ (number of instruments), $\rho = 0.4$ (correlation of u and v_2), $\mu^2 = 30$ (concentration parameter in (21)), homoskedastic case

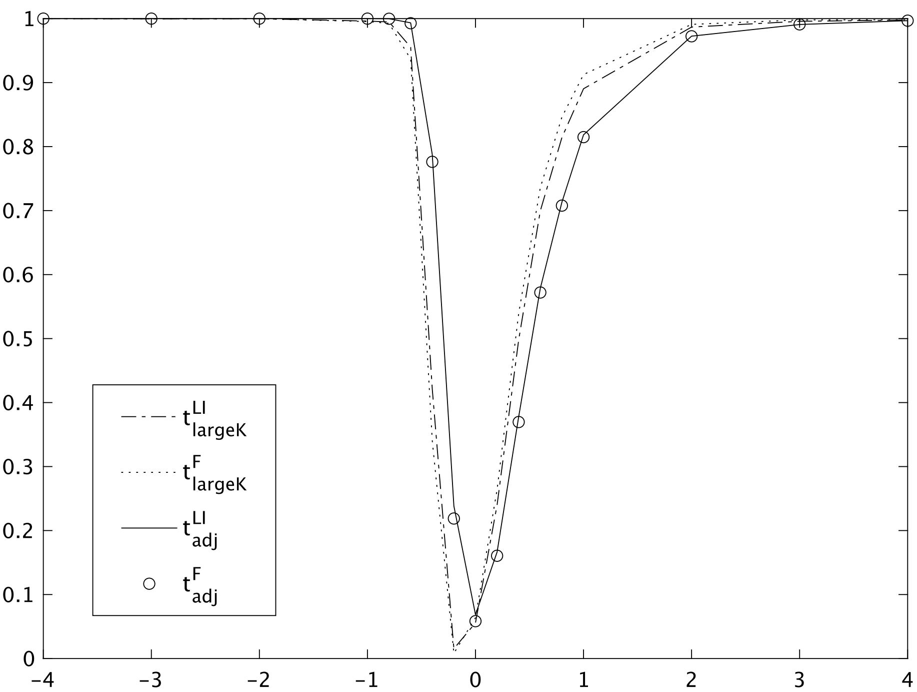

FIGURE 2. Power curves of two-sided tests: $n = 200$ (sample size), $K = 20$ (number of instruments), $\rho = 0.8$ (correlation of u and v_2), $\mu^2 = 30$ (concentration parameter in (21)), homoskedastic case

			t_K^H		$t_K^{H,F}$		$t_K^{H,adj}$	
ρ	μ^2	K	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$	$\beta < 0$	$\beta > 0$
0.4	60	5	0.037	0.068	0.029	0.071	0.055	0.061
	60	10	0.035	0.066	0.027	0.069	0.054	0.059
	60	20	0.033	0.067	0.025	0.070	0.053	0.061
	30	5	0.023	0.072	0.013	0.075	0.049	0.065
	30	10	0.023	0.072	0.013	0.074	0.048	0.067
	30	20	0.022	0.073	0.012	0.075	0.046	0.069
0.8	60	5	0.021	0.081	0.010	0.092	0.046	0.056
	60	10	0.021	0.080	0.009	0.092	0.046	0.056
	60	20	0.020	0.080	0.010	0.093	0.045	0.055
	30	5	0.005	0.092	0.000	0.106	0.026	0.060
	30	10	0.004	0.092	0.000	0.106	0.025	0.060
	30	20	0.003	0.094	0.000	0.110	0.020	0.062

Table 3. Null rejection frequencies of one-sided large-K t-tests at 5% significance level (heteroskedastic case). $t_K^{H,F}$ refers to the t-ratio based on the HFUL estimator by Hausman *et al.* (2012)

ρ	μ^2	K	t_K^H	$t_K^{H,F}$	$t_K^{H,adj}$
0.4	60	5	0.054	0.052	0.062
	60	10	0.051	0.049	0.059
	60	20	0.053	0.050	0.061
	30	5	0.048	0.046	0.058
	30	10	0.049	0.047	0.060
	30	20	0.049	0.046	0.061
0.8	60	5	0.060	0.060	0.050
	60	10	0.056	0.060	0.050
	60	20	0.055	0.059	0.049
	30	5	0.062	0.071	0.039
	30	10	0.062	0.071	0.039
	30	20	0.064	0.073	0.039

Table 4. Null rejection frequencies of two-sided large-K t-tests at 5% significance level (heteroskedastic case). $t_K^{H,F}$ refers to the t-ratio based on the HFUL estimator by Hausman *et al.* (2012)

References

[1] Anderson, T. W. (1974) An asymptotic expansion of the distribution of the limited information maximum likelihood estimate of a coefficient in a simultaneous equation system, *Journal of the American Statistical Association*, 69, 565-573.

[2] Anderson, T. W. (2003) *An Introduction to Multivariate Statistical Analysis*, 3rd edition, John-Wiley.

[3] Anderson, T. W., Kunitomo, N. and Y. Matsushita (2010) On the asymptotic optimality of the LIML estimator with possibly many instruments, *Journal of Econometrics*, 157, 191-204.

[4] Anderson, T. W., Kunitomo, N. and Y. Matsushita (2011) On finite sample properties of alternative estimators of coefficients in a structural equation with many instruments, *Journal of Econometrics*, 165, 58-69.

[5] Anderson, T. W., Kunitomo, N. and K. Morimune (1986) Comparing single-equation estimators in a simultaneous equations system, *Econometric Theory*, 2, 1-32.

[6] Anderson, T. W., Kunitomo, N. and T. Sawa (1982) Evaluation of the distribution function of the limited information maximum likelihood estimator, *Econometrica*, 50, 1009-1027.

[7] Angrist, J. D. and A. Krueger (1991) Does compulsory school attendance affect schooling and earnings, *Quarterly Journal of Economics*, 106, 979-1014.

[8] Bekker, P. A. (1994) Alternative approximations to the distributions of instrumental variables estimators, *Econometrica*, 63, 657-681.

[9] Chao, J. C., Swanson, N. R., Hausman, J. A., Newey, W. K. and T. Woutersen (2012) Asymptotic distribution of JIVE in a heteroskedastic IV regression with many instruments, *Econometric Theory*, 28, 42-86.

[10] Chioda, L. and M. Jansson (2009) Optimal invariant inference when the number of instruments is large, *Econometric Theory*, 25, 793-805.

[11] Fujikoshi, Y., Morimune, K., Kunitomo, N. and M. Taniguchi (1982) Asymptotic expansions of the distributions of the estimates of coefficients in a simultaneous equation system, *Journal of Econometrics*, 18, 191-205.

[12] Fuller, W. A. (1977) Some properties of a modification of the limited information estimator, *Econometrica*, 45, 939-954.

[13] Hahn, J., Hausman, J. A. and G. M. Kuersteiner (2004) Estimation with weak instruments: accuracy of higher-order bias and MSE approximations, *Econometrics Journal*, 7, 272-306.

[14] Hansen, C., Hausman, J. A. and W. K. Newey (2008) Estimation with many instrumental variables, *Journal of Business & Economic Statistics*, 26, 398-422.

[15] Hausman, J. A., Newey, W. K., Woutersen, T., Chao, J. and N. R. Swanson (2012) Instrumental variable estimation with heteroskedasticity and many instruments, *Quantitative Economics*, 3, 211-255.

[16] Kunitomo, N. (1980) Asymptotic expansions of distributions of estimators in a linear functional relationship and simultaneous equations, *Journal of the American Statistical Association*, 75, 693-700.

[17] Kunitomo, N. (1982) Asymptotic efficiency and higher order efficiency of the limited information maximum likelihood estimator in large econometric models, Technical Report No. 365, Institute for Mathematical Studies in the Social Sciences, Stanford University.

[18] Kunitomo, N. and Y. Matsushita (2009) Asymptotic expansions of the distributions of semi-parametric estimators in a linear simultaneous equations system, *Journal of Multivariate Analysis*, 100, 1727-1751.

[19] Morimune, K. (1983) Approximate distributions of k-class estimators when the degree of overidentification is large compared with sample size, *Econometrica*, 51, 821-841.

[20] Morimune, K. (1989) t test in a structural equation, *Econometrica*, 57, 1341-1360.

[21] Mukerjee, R. (1992) Comparison between the conditional likelihood ratio and the usual likelihood ratio tests, *Journal of the Royal Statistical Society*, B 54, 189-194.

[22] Phillips, P. C. B. (1977) A general theorem in the theory of asymptotic expansions as approximations to the finite sample distributions of econometric estimators, *Econometrica*, 45, 1517-1534.

[23] Rao, C. R. (1973) *Linear Statistical Inference and Its Applications*, 2nd edition, John-Wiley.

[24] Rothenberg, T. J. (1984) Approximating the distributions of econometric estimators and test statistics, in Griliches, Z. and M. D. Intriligator (eds.) *Handbook of Econometrics*, vol. II, ch. 15, Elsevier.

[25] Rothenberg, T. J. (1988) Approximate power functions for some robust tests of regression coefficients, *Econometrica*, 56, 997-1019.

[26] Sargan, J. D. (1975) Gram-Charlier approximations applied to t ratios of k-class estimators, *Econometrica*, 43, 327-346.

[27] Ullah, A. (2004) *Finite Sample Econometrics*, Oxford University Press.

[28] van Hasselt, M. (2010) Many instruments asymptotic approximations under non-normal error distributions, *Econometric Theory*, 26, 633-645.

Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan.

Email address: matsushita.y@r.hit-u.ac.jp

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.

Email address: t.otsu@lse.ac.uk